



# Stability analysis of the Interior Penalty Discontinuous Galerkin method for the wave equation

Cyril Agut, Julien Diaz

## ► To cite this version:

Cyril Agut, Julien Diaz. Stability analysis of the Interior Penalty Discontinuous Galerkin method for the wave equation. ESAIM: Mathematical Modelling and Numerical Analysis, 2013, 47 (3), pp.903-932. 10.1051/m2an/2012061 . hal-00759457

**HAL Id: hal-00759457**

**<https://inria.hal.science/hal-00759457>**

Submitted on 21 Nov 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# STABILITY ANALYSIS OF THE INTERIOR PENALTY DISCONTINUOUS GALERKIN METHOD FOR THE WAVE EQUATION

CYRIL AGUT<sup>1</sup> AND JULIEN DIAZ<sup>2</sup>

**Abstract.** We consider here the Interior Penalty Discontinuous Galerkin (IPDG) discretization of the wave equation. We show how to derive the optimal penalization parameter involved in this method in the case of regular meshes. Moreover, we provide necessary stability conditions of the global scheme when IPDG is coupled with the classical Leap-Frog scheme for the time discretization. Numerical experiments illustrate the fact that these conditions are also sufficient.

**1991 Mathematics Subject Classification.** 35L05, 65M12, 65M60.

## 1. INTRODUCTION

Nowadays, in many domains as for instance medical imaging, seismic imaging, radar, earthquakes simulations, one requires the accurate numerical solution to acoustic, elastodynamic or electromagnetic wave equations. In many cases, the equations have to be solved in very large domains with strong heterogeneities. Therefore, one has to use sophisticated discretization methods to compute the most accurate solution for the smallest computational cost.

Finite Differences Methods, which are widely used because of their small computational cost and the simplicity of their implementation, are not adapted to deal with strong heterogeneities. Indeed, they rely on structured grids, which can not accurately approximate the shape of the various layers of the domain. Finite Elements Methods (FEM) are more adapted to this kind of problems since they allow for the use of unstructured grids. Among all existing FEM, the Spectral Element Method [8, 10, 15, 16], is probably one of the most efficient to solve the wave equation since the resulting mass matrix is diagonal. Hence, it can be easily coupled to explicit time-schemes, which requires the inversion of the mass matrix at each time step. However, SEM requires the use of quadrilateral (in 2D) or hexahedral (in 3D) meshes which can be difficult to generate for realistic applications. Let us however mention that SEM can be extended to handle triangular meshes [9], but it requires additional degrees of freedom and the implementation is more complex. Moreover, as far as we know, the extension to tetrahedral meshes has not been proposed yet.

Discontinuous Galerkin Methods are more and more popular for solving the wave equation since they lead to block-diagonal mass matrices without the help of quadrature formula. Moreover, they can be used with any type of meshes and even allow for the variation of the physical parameters inside the cells of the mesh (provided that the variation is polynomial). DGM are also naturally adapted to parallel computing since all

---

*Keywords and phrases:* Discontinuous Galerkin, penalization coefficient, CFL condition, wave equation

<sup>1</sup> LMAP - University of Pau - INRIA Project-Team Magique-3D

<sup>2</sup> INRIA Project-Team Magique-3D - LMAP - University of Pau

volume integrals are computed locally and the communications between the cells are ensured by integrals over the faces of the elements. In [4], the authors provide a detailed review of the various Discontinuous Galerkin approximations of the Laplacian operator. They show that the so-called Interior Penalty Discontinuous Galerkin Method (IPDGM), also known as Symmetric Interior Penalty (SIP) [3,6], is one of the most suitable since it is stable and adjoint consistent, which guarantees the optimal order of convergence of the scheme. This explains why this method has been successfully used to solve Helmholtz equation [2, 7] and the wave equation [2, 13, 14]. Comparisons of the performances of IPDGM and SEM can be found in [11] and [5]. It is worth noting that the first paper concluded that the performances of SEM are better than IDPGM when IPDGM is applied to structured grids composed of squares, while the second paper show that the performances of IPDGM are better when it is applied to triangular meshes.

Nevertheless, in spite of all its interesting properties, IPDGM still suffers from two difficulties. The first one is the determination of the penalization parameter, which penalizes the discontinuities of the solution through the faces. The accurate determination of the optimal parameter is crucial, since a too small value leads to instabilities while a too large value could (strongly) hamper the CFL (Courant-Friedrichs-Lewy) condition, which gives the maximal time step that can be used to ensure the stability of the scheme. In [2], the authors conjectured a minimal value of the penalization parameters, depending on  $p$ , the polynomial degree of the basis functions and on the size of the elements. They proved their conjecture up to  $p = 3$ . The extension of this result to  $p > 3$  and to unstructured meshes, is still to be done. The second difficulty is the determination of the CFL condition. It is well-known that this condition decreases when the penalization parameter increases, but no analytical formula has been proposed yet. The aim of this paper is a) to prove the conjecture of Ainsworth, Monk and Muniz up to  $p = 5$ ; b) to propose a solution methodology to prove it for a given  $p$ ; and c) to provide an analytic formula linking the CFL condition to the penalization parameter. We restrict ourselves to the cases of structured meshes composed of segments (in 1D), squares (in 2D) or cubes (in 3D). In section 1, we recall the IPDG discretization of the wave equation. In section 2, we propose two theorems, the first one provides explicit necessary stability conditions on the penalization parameter and the time step while the second one provides a more restrictive but implicit necessary stability condition. The proof of these theorems in the one dimensional case is given in section 3. Section 4 is devoted to the proof in the three dimensional case and contains a discussion on the adaptation of the theorems to structured meshes composed of rectangles or parallelepipeds. We do not present the proof in the two dimensional case, but it can be adapted without any difficulties from the three dimensional case. Finally, we present numerical results in section 5 that illustrate the fact that the stability condition is actually necessary and sufficient for numerical applications.

## 2. INTERIOR PENALTY DISCONTINUOUS GALERKIN DISCRETIZATION OF THE WAVE EQUATION

In this section, we recall the so called Interior Penalty Discontinuous Galerkin method applied to the acoustic wave equation in homogeneous bounded media  $\Omega \subset \mathbb{R}^d, d = 1, 2, 3$ . For the sake of simplicity, we impose homogeneous Dirichlet boundary conditions on the boundary  $\Gamma := \partial\Omega$  but this study can be extended to Neumann boundary conditions without major difficulties.

$$\left\{ \begin{array}{ll} \text{Find } u : \Omega \times [0, T] \mapsto \mathbb{R} \text{ such that :} & \\ \frac{1}{\mu} \frac{\partial^2 u}{\partial t^2} - \operatorname{div} \left( \frac{1}{\rho} \nabla u \right) = f & \text{in } \Omega \times ]0, T], \\ u(x, 0) = u_0, \quad \frac{\partial u}{\partial t}(x, 0) = u_1 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{array} \right. \quad (1)$$

where  $u$  stands for the displacement,  $\mu$  is the compressibility modulus,  $\rho$  is the density and  $f$  is the source term. We introduce a triangulation  $\mathcal{T}_h$  of  $\Omega$  and the following space of approximation with piecewise discontinuous polynomial functions :

$$V_h := \{v \in L^2(\Omega) : v|_K \in P^p(K), \forall K \in \mathcal{T}_h\}.$$

The set of the mesh faces is denoted  $\mathcal{F}_h$  which is partitionned into two subsets  $\mathcal{F}_h^i$  and  $\mathcal{F}_h^b$  corresponding respectively to the interior faces and those located on the boundary. For  $F \in \mathcal{F}_h^i$ , we note arbitrarily  $K^+$  and  $K^-$  the two elements sharing  $F$  and we define  $\nu$  as the unit outward normal vector pointing from  $K^+$  to  $K^-$ . Moreover,  $v^\pm$  represents the restriction of a function  $v$  to the element  $K^\pm$  and we define the jump and the average of a piecewise smooth function  $v \in V_h$  over  $F \in \mathcal{F}_h^i$  such that

$$[[v]] = v^+ - v^-, \quad \{\{v\}\} = \frac{v^+ + v^-}{2}. \quad (2)$$

For  $F \in \mathcal{F}_h^b$ , we define  $[[v]] = v$  and  $\{\{v\}\} = v$ . The IPDG discretization of (1) reads as

$$\begin{cases} \text{Find } u_h \in V_h \text{ such that, } & \forall v_h \in V_h : \\ \sum_{K \in \mathcal{T}_h} \int_K \frac{1}{\mu} \partial_t^2 u_h v_h dx = & -a_h(u_h, v_h) + \sum_{K \in \mathcal{T}_h} \int_K f v_h dx. \end{cases} \quad (3)$$

where  $a_h$  is a bilinear form defined by

$$a_h(u_h, v_h) = B_{\mathcal{T}_h}(u_h, v_h) - \mathcal{I}(u_h, v_h) - \mathcal{I}(v_h, u_h) + B_S(u_h, v_h),$$

with

$$B_{\mathcal{T}_h}(u_h, v) = \sum_{K \in \mathcal{T}_h} \int_K \frac{1}{\rho} \nabla u_h \nabla v, \quad \mathcal{I}(u_h, v) = \sum_{F \in \mathcal{F}_h^i} \int_F [[v]] \left\{ \left\{ \frac{1}{\rho} \nabla u_h \cdot \nu \right\} \right\},$$

$$B_S(u_h, v) = \sum_{F \in \mathcal{F}_h^i} \int_F \gamma [[u_h]] [[v]].$$

The bilinear form  $B_S$  is devoted to enforce the coercivity of  $a_h$  and the penalization function  $\gamma$  is defined on each interior face  $F$  by

$$\gamma = \frac{\alpha}{\xi_F}$$

where  $\alpha$  is a positive parameter. There are many definitions of the function  $\xi_F$  in the litterature. The most commonly used are:

- $\xi_F = h(F)$  where  $h(F)$  denotes the diameter of  $F$ . See for instance [2, 4, 14]. It is worth noting that this definition does not make sense in 1D.
- $\xi_F = \min(h(K^+), h(K^-))$  where  $h(K^\pm)$  is the diameter of  $K^\pm$ . See for instance [13].
- $\xi_F = \min(\rho(K^+), \rho(K^-))$  where  $\rho(K^\pm)$  is the diameter of the inscribed circle (or sphere) of  $K^\pm$ . See for instance [17].

Whatever the definition of  $\xi_F$ , there exists  $\alpha_0^p > 0$  such that the coercivity of  $a_h$  is ensured for all  $\alpha \geq \alpha_0^p$ . Obviously, the optimal parameter  $\alpha_0$  depends on the choice of the basis functions of  $V_h$ , but also on  $\xi_F$ . It has been shown by Shabazi in [17], using inverse inequalities proposed in [18], that the third definition was the most appropriate for triangular meshes. In [2], the authors showed that, on square and cubic meshes, using the first

definition of  $\xi_F$ , the optimal parameter is  $\alpha_0^p = \frac{p(p+1)}{2}$  for  $p = 0, \dots, 3$  and they conjectured this relation for  $p \geq 4$ . It is worth noting that, for such meshes, the first and the third definition of  $\xi_F$  are equivalent.

At this point, we choose not to explicit the expression of  $\xi_F$ . This will be done in the next section.

We refer to [2, 4, 13] for more details on the properties of the bilinear form  $a_h$ .

Considering  $\{\varphi_i\}_{i=1,\dots,m}$  the classical discontinuous Lagrange basis functions of degree  $p$  of  $V_h$ , where  $m$  denotes the number of degrees of freedom of the problem, we obtain the following linear system:

$$\partial_t^2 U = M^{-1} K U + M^{-1} F \quad (4)$$

where

$$(M)_{i,j} = \sum_{K \in \mathcal{T}_h} \int_K \varphi_i \varphi_j, \quad (K)_{i,j} = a_h(\varphi_i, \varphi_j), \quad (F)_i = \sum_{K \in \mathcal{T}_h} \int_K f \varphi_i.$$

Now, we have to discretize in time. Using the well known Leap-Frog scheme, we obtain the following fully discretized scheme:

$$\frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} = -M^{-1} K U^n + M^{-1} F^n. \quad (5)$$

Since (5) is an explicit scheme, its  $L^2$  stability is constrained by a CFL condition. It is well known [14] that this CFL condition decreases when  $\alpha$  increases and behaves as  $\frac{1}{\sqrt{\alpha}}$  for large  $\alpha$ . However, no explicit formula of the CFL condition has been proposed yet. This is the objective of the next section.

### 3. STABILITY ANALYSIS

In this section, we first propose necessary conditions over  $\gamma$  and  $\Delta t$  ensuring the  $L^2$ -stability of scheme (5). This theorem provides an explicit dependence of  $\Delta t$  with respect to  $\gamma$  and  $h$ . Next we propose a more restrictive necessary stability condition. In this second theorem, the dependence of  $\Delta t$  with respect to  $\gamma$  is no longer explicit. However the condition can be numerically computed using the roots of a polynomial of degree  $2p$ . We assume here that the domain  $\Omega$  is unbounded ( $\Omega = \mathbb{R}^d$ ) and uniformly meshed by segments (if  $d = 1$ ), squares (if  $d = 2$ ) or cubes (if  $d = 3$ ). The length of the edges of the elements is denoted by  $h$ .

The necessary stability conditions are given by the following theorem.

**Theorem 3.1.** *The scheme (5) is  $L^2$ -stable only if, for  $p \leq 5$ ,*

$$\gamma \geq \frac{p(p+1)}{2h} = \frac{\alpha_p^0}{h}; \quad (6)$$

and, denoting  $\gamma = \alpha/h$ ,

$$\sqrt{d} \frac{c \Delta t}{h} \leq \begin{cases} C_{1,p} & \text{if } \alpha \leq \alpha_{1,p} \\ C_{2,p}(\alpha) & \text{if } \alpha > \alpha_{1,p}. \end{cases} \quad (7)$$

where  $\alpha_{1,p}$ ,  $C_{1,p}$  and  $C_{2,p}(\alpha)$  are defined with respect to the polynomial degree  $p$  such that:

$p$	$\alpha_{1,p}$	$C_{1,p}$	$C_{2,p}(\alpha)$
1	2	$\frac{\sqrt{3}}{3} \simeq 0.577$	$\frac{1}{\sqrt{3(\alpha-1)}}$
2	$\frac{27}{5} = 5.4$	$\frac{1}{\sqrt{15}} \simeq 0.258$	$\sqrt{\frac{2}{-15 + 6\alpha + (405 - 240\alpha + 36\alpha^2)^{1/2}}}$
3	$\frac{2\sqrt{1605} + 393}{49} \simeq 9.65$	$\sqrt{\frac{2}{45 + \sqrt{1605}}} \simeq 0.153$	$\sqrt{\frac{2}{-45 + 10\alpha + (4545 - 1320\alpha + 100\alpha^2)^{1/2}}}$
4	$\alpha_{1,4} \simeq 14.7$	$\sqrt{\frac{2}{3(35 + \sqrt{805})}} \simeq 0.103$	$\sqrt{\frac{1}{2\sqrt{5}g_{4,1}(\alpha)g_{4,2}(\alpha) + 5\alpha - 35}}$
5	$\alpha_{1,5} \simeq 20.8$	$\sqrt{\frac{1}{10\sqrt{133}\cos(g_5) + 70}} \simeq 0.074$	$\sqrt{\frac{1}{2\sqrt{7}g_{5,1}(\alpha)g_{5,2}(\alpha) + 7(\alpha - 10)}}$

where, for the case  $p = 4$ , we have

$$\begin{cases} g_{4,1}(\alpha) = (518 - 98\alpha + 5\alpha^2)^{\frac{1}{2}}, \\ g_{4,2}(\alpha) = \cos\left(\frac{1}{3}\arccos\left(\frac{1}{10}g_{4,3}(\alpha)\frac{\sqrt{5}}{g_{4,1}^3(\alpha)}\right)\right), \\ g_{4,3}(\alpha) = -47705 + 14574\alpha - 1470\alpha^2 + 50\alpha^3 \end{cases}$$

and for the polynomials of degree 5,

$$\begin{cases} g_5 = \frac{1}{3}\arccos\left(\frac{10447}{126350}\sqrt{133}\right) \\ g_{5,1}(\alpha) = (1555 - 200\alpha + 7\alpha^2)^{\frac{1}{2}}, \\ g_{5,2}(\alpha) = \cos\left(\frac{1}{3}\arccos\left(\frac{1}{14}g_{5,3}(\alpha)\frac{\sqrt{7}}{g_{5,1}^3(\alpha)}\right)\right), \\ g_{5,3}(\alpha) = -299825 + 61440\alpha - 4200\alpha^2 + 98\alpha^3. \end{cases}$$

**Remark 3.2.**

- As it is was noted in [14], the stability condition on  $\Delta t$  behaves as  $C/\sqrt{\alpha}$  for  $\alpha$  large enough. More precisely,  $C = \sqrt{\frac{2}{(p+1)(p+2)}}$ .
- This stability condition is constant for  $\frac{p(p+1)}{2} \leq \alpha \leq \alpha_p^1$ . This shows that it is not necessary to choose  $\alpha$  too close from  $\alpha_p^0$  to improve the CFL condition.
- In [2], they authors proved (6) for  $p = 0, \dots, 3$  and conjectured this relation for any  $p$ . Theorem (3.3) extends its validity until  $p = 5$ .
- The condition (6) does not depend on the dimension  $d$ . This would not have been the case if we had expressed  $\gamma$  as a function of the circumcircle (or circumsphere) diameter which is  $\sqrt{d}h$ . Since  $h$  is

the diameter of the inscribed circle or sphere, we conjecture that the third definition of  $\xi_F$  is the most appropriate. We will strengthen this conjecture when we discuss the extension of this theorem to meshes composed of rectangles or parallelepipeds.

We have performed numerical experiments to compute the numerical CFL condition on finite meshes. We have observed (see section 6) that this condition was equivalent to condition (7), except for a small range of value of  $\alpha$ . Therefore we need a more restrictive necessary condition provided by the following theorem.

**Theorem 3.3.** *Let  $V_{p,\alpha} = \left\{ \lambda \in \mathbb{R} : Q_{p,\alpha}(\lambda) = 0 \text{ and } |\tilde{Q}_{p,\alpha}(\lambda)| \leq 1 \right\}$  where  $Q_{p,\alpha}(\lambda)$  is a polynomial of degree  $2p$  and  $\tilde{Q}_{p,\alpha}(\lambda)$  is a rational function whose expressions are given in appendix B. Then, scheme (5) is  $L^2$ -stable only if, for  $p \leq 5$ , the conditions (6) and (7) are satisfied and*

$$\begin{cases} V_{p,\alpha} = \emptyset \\ \text{or} \\ \sqrt{d} \frac{c\Delta t}{h} \leq C_{3,p}(\alpha) = \frac{1}{2\sqrt{\min V_{p,\alpha}}}. \end{cases} \quad (8)$$

**Remark 3.4.**

- This theorem does not provide an explicit CFL condition. However, it can be computed numerically by the following algorithm:
  - (1) Compute all the roots of  $Q_{p,\alpha}$ ,
  - (2) Select the real roots such that  $|\tilde{Q}_{p,\alpha}(\lambda)| \leq 1$ ,
  - (3) Choose the maximum of these roots.
- The numerical results in section 6 show that this theorem gives in practical cases necessary and sufficient conditions.
- The numerical study of condition (8) that we present in section 6 shows that the set  $V_{p,\alpha}$  is actually empty except when  $\alpha$  belongs to a small segment around  $\alpha_1^p$ . This means that theorem 3.1 provides a sufficient and necessary stability condition when  $\alpha$  is not in this segment. Moreover the remarks 3.2 are still valid.

We were unfortunately unable to establish this theorem for any  $p$  and we have restricted ourselves to  $p \leq 5$ . The proofs in the one dimensional case are given in section 4 while the extension to  $d = 3$  is the subject of section 5. The proof for  $d = 2$  can be easily deduced from the case  $d = 3$ .

#### 4. STUDY IN THE 1-DIMENSIONAL CASE

This section contains the proofs of theorems 3.1 and 3.3 in the one dimensional case. It consists of three steps. The first step is a Fourier analysis presented in section 4.1; the second step is devoted to the proof of conditions (6) and is presented in section 4.2; the last step concerns the proof of (7) and (8) in section 4.3. The proofs are detailed for  $p = 3$  and are easily extendable to the cases  $p = 1, 2, 4$  and 5.

Here, we assume that the domain is  $\Omega = \mathbb{R}$  and is meshed by segments of length  $h$ . We consider a velocity  $c^2 = \mu/\rho = 1$  but we can extend the proof to other velocities by setting  $\Delta t' = \Delta t/c$ . We consider the scheme (5) without source term that is to say

$$M \frac{U^{n+1} - 2U^n + U^{n-1}}{\Delta t^2} + KU^n = 0. \quad (9)$$

Considering the equation on one element  $J$  of the mesh, we have  $\forall J \in \mathcal{T}_h$

$$M_{1,p} \frac{U_J^{n+1} - 2U_J^n + U_J^{n-1}}{\Delta t^2} + (K_{1,p}^W)^T U_{J-1}^n + K_{1,p} U_J^n + K_{1,p}^W U_{J+1}^n = 0 \quad (10)$$

where  $U_J$  corresponds to the vector of unknowns  $U$  restricted to the element  $J$  and  $M_{1,p}$ ,  $K_{1,p}$  and  $K_{1,p}^W$  are respectively the mass and stiffness matrices in dimension 1 considering polynomials of degree  $p$ :

$$\begin{aligned}
M_{1,p}(i, j) &= h \int_{[0,1]} \hat{\varphi}_i(\hat{x}) \hat{\varphi}_j(\hat{x}) d\hat{x}, \\
K_{1,p}(i, j) &= \frac{1}{h} \int_{[0,1]} \frac{\partial \hat{\varphi}_i}{\partial \hat{x}}(\hat{x}) \frac{\partial \hat{\varphi}_j}{\partial \hat{x}}(\hat{x}) d\hat{x} + \frac{1}{2h} \hat{\varphi}_i(1) \frac{\partial \hat{\varphi}_j}{\partial \hat{x}}(1) + \frac{1}{2h} \hat{\varphi}_j(1) \frac{\partial \hat{\varphi}_i}{\partial \hat{x}}(1) \\
&\quad + \gamma \hat{\varphi}_i(1) \hat{\varphi}_j(1) - \frac{1}{2h} \hat{\varphi}_i(0) \frac{\partial \hat{\varphi}_j}{\partial \hat{x}}(0) - \frac{1}{2h} \hat{\varphi}_j(0) \frac{\partial \hat{\varphi}_i}{\partial \hat{x}}(0) \\
&\quad + \gamma \hat{\varphi}_i(0) \hat{\varphi}_j(0), \\
K_{1,p}^W(i, j) &= -\frac{1}{2h} \hat{\varphi}_i(1) \frac{\partial \hat{\varphi}_j}{\partial \hat{x}}(0) + \frac{1}{2h} \hat{\varphi}_j(0) \frac{\partial \hat{\varphi}_i}{\partial \hat{x}}(1) - \gamma \hat{\varphi}_i(1) \hat{\varphi}_j(0),
\end{aligned} \tag{11}$$

where  $\{\hat{\varphi}_i\}_{i=1,\dots,p+1}$  are the classical discontinuous Lagrange basis functions on the reference element  $[0, 1]$ .

#### 4.1. Fourier Analysis of the IPDG scheme in 1D

In order to study the stability of the IPDG scheme, we have to introduce the discrete Fourier transform

$$\begin{aligned}
\mathcal{F}_h : L_h^2 &\rightarrow L^2(K_h) \\
U &\rightarrow \tilde{U} = \mathcal{F}_h(U)(k) = \frac{h}{2\pi} \sum_{J \in \mathbb{Z}} U_J e^{-ikJh}
\end{aligned}$$

with  $K_h = [-\frac{\pi}{h}, \frac{\pi}{h}]$  and  $L_h^2 = \left\{ U = (U_J)_{J \in \mathbb{Z}}, \sum_{J \in \mathbb{Z}} \|U_J\|^2 < +\infty \right\}$ .

Now, applying this discrete Fourier transform to (10), we obtain,  $\forall \beta \in [-\pi, \pi]$

$$M_{1,p} \frac{\tilde{U}_J^{n+1}(\beta) - 2\tilde{U}_J^n(\beta) + \tilde{U}_J^{n-1}(\beta)}{\Delta t^2} + K_\beta \tilde{U}_J^n(\beta) = 0 \tag{12}$$

where  $\beta = hk$  and  $K_\beta = (K_{1,p}^W)^T e^{-i\beta} + K_{1,p} + K_{1,p}^W e^{i\beta}$ .

The  $L^2$ -stability of (12) for all  $\beta \in [-\pi, \pi]$ , is equivalent to the  $L^2$ -stability of (9), thanks to the Parseval equalities.

Since  $M_{1,p}$  is positive definite and  $K_\beta$  is hermitian, all the eigenvalues of  $N_\beta = M_{1,p}^{-1} K_\beta$  are real. Moreover, a classical stability analysis shows that (12) is stable if and only if

$$0 \leq \lambda \leq \frac{4}{\Delta t^2}$$

for all  $\lambda \in \Lambda(\beta)$ , where  $\Lambda(\beta)$  denotes the set of the eigenvalues of  $N_\beta$ . Then, a necessary and sufficient condition of the stability of (9) is

$$\lambda_{\min} \geq 0 \quad \text{and} \quad \Delta t \leq \frac{2}{\sqrt{\lambda_{\max}}}$$

with  $\lambda_{\min} = \min_{\beta \in [-\pi, \pi]} [\min(\Lambda(\beta))]$  and  $\lambda_{\max} = \max_{\beta \in [-\pi, \pi]} [\max(\Lambda(\beta))]$ .

In section 4.2, we show that the condition  $\lambda_{\min} \geq 0$  is equivalent to (6) and in section 4.3, we show that the condition  $\lambda_{\max} \leq \frac{4}{\Delta t^2}$  implies (7) and (8).



#### 4.2. Study of the condition $\lambda_{\min} \geq 0$

In the following, we consider the change of variable  $\alpha = h\gamma$  to simplify the presentation. To show the equivalence between (3.1) and  $\lambda_{\min} \geq 0$ , we have to consider the characteristic polynomial of  $N_\beta$ :

$$q_\alpha(\lambda, \beta) = (-1)^{p+1} \lambda^{p+1} + \sum_{i=0}^p c_i(\alpha, \beta) \lambda^i.$$

The coefficients  $c_i(\alpha, \beta)$  can be computed by a symbolic calculus software such as Maple. We present them in appendix A for  $1 \leq p \leq 5$ .

In order to study the sign of the eigenvalues of  $N_\beta$ , we will use the following lemma.

**Lemma 4.1.** *Let  $P$  be a polynomial of degree  $n$  with  $n$  real roots such that  $P(X) = \sum_{i=0}^n c_i X^i$ . All the roots of  $P$  are non negative if and only if*

$$(-1)^i c_i \geq 0.$$

*Proof.* The proof of this lemma can be found in [1]. □

Hence, we have to find a condition over  $\alpha$  such that,  $\forall i \in \{0, \dots, p\}$ ,  $\forall \beta \in [-\pi, \pi]$ ,

$$(-1)^i c_i(\alpha, \beta) \geq 0.$$

Here we only detail the computations in the case  $p = 3$  and we give the expression of the characteristic polynomials for  $p \neq 3$  in the appendix A. For  $p = 3$ , we have

$$\begin{cases} c_3(\alpha, \beta) = \frac{8}{h^2} ((15 - \alpha) \cos(\beta) - 4\alpha) \\ c_2(\alpha, \beta) = \frac{240}{h^4} (\cos^2(\beta) - (23 + \alpha) \cos(\beta) + (18\alpha - 65)) \\ c_1(\alpha, \beta) = \frac{2880}{h^6} (4 \cos^2(\beta) + (65 - 3\alpha) \cos(\beta) + (141 - 32\alpha)) \\ c_0(\alpha, \beta) = \frac{100800}{h^8} (3 \cos^2(\beta) + 2(3 - \alpha) \cos(\beta) + (2\alpha - 9)). \end{cases}$$

- Let us first study the condition over  $c_3$ . We have,  $\forall \beta \in [-\pi, \pi]$ ,

$$-c_3(\alpha, \beta) \geq 0 \Leftrightarrow (\alpha - 15) \cos(\beta) + 4\alpha \geq 0.$$

It is clear that this condition is satisfied for all  $\beta$  if and only if

$$\begin{cases} (\alpha - 15) + 4\alpha \geq 0, \\ (15 - \alpha) + 4\alpha \geq 0 \end{cases} \quad (13)$$

which implies that

$$\begin{cases} \alpha \geq 3, \\ \alpha \geq -5. \end{cases} \quad (14)$$

Consequently,  $-c_3(\alpha, \beta) \geq 0$ , whatever the choice of  $\beta$ , if and only if

$$\alpha \geq 3 \quad (15)$$

- Let us now consider the condition over  $c_2$ .  
Setting  $X = \cos(\beta)$ , this condition is equivalent to

$$f_\alpha(X) := X^2 - (23 + \alpha)X + (18\alpha - 65) \geq 0, \forall X \in [-1, 1]. \quad (16)$$

This second order polynomial admits two roots:

$$\begin{cases} X_1 = \frac{1}{2} \left( 23 + \alpha + \left( (\alpha - 13)^2 + 620 \right)^{1/2} \right), \\ X_2 = \frac{1}{2} \left( 23 + \alpha - \left( (\alpha - 13)^2 + 620 \right)^{1/2} \right). \end{cases}$$

We know that  $f_\alpha(X)$  is a second-order polynomial on the variable  $X$  and its head coefficient is positive. Thus, to have  $f_\alpha$  non negative  $\forall X \in [-1, 1]$ , we need one of the following conditions:

- (1) the two roots are in  $] -\infty; -1]$  i.e.  $X_1 \leq -1$ ,
- (2) the two roots are in  $]1; +\infty[$  i.e.  $X_2 \geq 1$ ,
- (3)  $X_1 = X_2$ .

Since  $(\alpha - 13)^2 + 620 > 0$ ,  $X_1 > X_2$  and 3. is impossible.

The case  $X_1 \leq -1$  is also impossible since  $X_1 \geq 0$  when  $\alpha \geq 0$  so we just have to consider the case  $X_2 \geq 1$ , which leads to the inequality

$$23 + \alpha - \left( (\alpha - 13)^2 + 620 \right)^{1/2} \geq 2,$$

which is equivalent to

$$\alpha \geq \frac{87}{17}. \quad (17)$$

Finally,  $c_2(\alpha, \beta) \geq 0$ , whatever the choice of  $\beta$ , if and only if (17) holds.

- Now, let us study the sign of  $c_1(\alpha, \beta)$ .  
Using the change of variable  $X = \cos(\beta)$ , the condition  $-c_1(\alpha, \beta) \geq 0, \forall \beta \in [-\pi, \pi]$  is equivalent to

$$f_\alpha(X) := -4X^2 + (3\alpha - 65)X + (32\alpha - 141) \geq 0, \forall X \in [-1, 1].$$

The polynomial  $f_\alpha$  admits the two following roots:

$$\begin{cases} X_1 = \frac{-1}{8} \left( 65 - 3\alpha + \left( \left( 3\alpha + \frac{61}{3} \right)^2 + \frac{14000}{9} \right)^{1/2} \right), \\ X_2 = \frac{-1}{8} \left( 65 - 3\alpha - \left( \left( 3\alpha + \frac{61}{3} \right)^2 + \frac{14000}{9} \right)^{1/2} \right). \end{cases}$$

Since, the head coefficient of the polynom  $f_\alpha$  is negative, we need  $X_1 \leq -1$  and  $X_2 \geq 1$ .

The condition  $X_1 \leq -1$  implies that

$$65 - 3\alpha + \left( \left( 3\alpha + \frac{61}{3} \right)^2 + \frac{14000}{9} \right)^{1/2} \geq 8$$

which leads to

$$\alpha \geq \frac{80}{29}. \quad (18)$$

In the same way, the condition  $X_2 \geq 1$  is equivalent to

$$\alpha \geq 6. \quad (19)$$

Consequently,  $-c_1(\alpha, \beta) \geq 0, \forall \beta \in [-\pi, \pi]$  if and only if  $\alpha \geq 6$ .

- Finally, let us look at the positivity of  $c_0(\alpha, \beta), \forall \beta \in [-\pi, \pi]$ .  
Once again, using the change of variable  $X = \cos(\beta)$ , we have

$$f_\alpha(X) := 3X^2 + 2(3 - \alpha)X + 2\alpha - 9 \geq 0.$$

This polynomial function  $f_\alpha$  admits the two following roots:

$$\begin{cases} X_1 = \frac{1}{3}(2\alpha - 9), \\ X_2 = 1. \end{cases}$$

In the same way than previously, we need  $X_1 \geq 1$  which leads to the condition

$$\alpha \geq 6. \quad (20)$$

In conclusion, taking into account the conditions (15), (17), (18), (19) and (20), we have

$$\lambda_{\min} \geq 0 \Leftrightarrow \alpha \geq 6. \quad (21)$$

We used the same technique to derive a condition over  $\alpha$  for all polynomial degrees  $p$  from  $p = 1$  to  $p = 5$ . Since the calculations are very similar, we do not detail them here and we just present the conditions in Tab. 1.

$p$	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
1	$\alpha \geq 1$	$\alpha \geq 1$				
2	$\alpha \geq 3$	$\alpha \geq \frac{30}{11}$	$\alpha \geq 2$			
3	$\alpha \geq 6$	$\alpha \geq 6$	$\alpha \geq \frac{87}{17}$	$\alpha \geq 3$		
4	$\alpha \geq 10$	$\alpha \geq \frac{543}{55}$	$\alpha \geq \frac{325}{34}$	$\alpha \geq \frac{581}{73}$	$\alpha \geq 4$	
5	$\alpha \geq 15$	$\alpha \geq 15$	$\alpha \geq \frac{336}{23}$	$\alpha \geq \frac{1185}{86}$	$\alpha \geq \frac{124}{11}$	$\alpha \geq 5$

TABLE 1. Conditions over  $\alpha$  for each coefficient  $c_i$  and each polynomial degree  $p$

From these results, we can easily deduce the smallest penalization parameters ensuring the stability of the scheme (see Tab. 2). It is clear that, for  $1 \leq p \leq 5$ , the stability is guaranteed if and only if

$$\alpha \geq \frac{p(p+1)}{2}, \text{ or, equivalently, } \gamma \geq \frac{p(p+1)}{2h}.$$

$p$	1	2	3	4	5
	$\alpha \geq 1$	$\alpha \geq 3$	$\alpha \geq 6$	$\alpha \geq 10$	$\alpha \geq 15$

TABLE 2. Stability condition over  $\alpha$  for each polynomial degree  $p$ 

### 4.3. The CFL condition

Now, we propose to prove that the condition

$$\Delta t \leq \frac{2}{\sqrt{\lambda_{\max}}}$$

implies (7) and (8). To compute  $\lambda_{\max}$ , we use the implicit function theorem. Indeed, each eigenvalue  $\lambda \in \Lambda(\beta)$  is an implicit function  $\lambda(\beta)$  defined by  $p_\alpha(\lambda(\beta), \beta) = 0$ . Since  $p_\alpha(\lambda, \beta)$  is a continuous and periodic function of  $\beta$ , so is  $\lambda(\beta)$  and we only need to find all the values  $\beta_0 \in [-\pi; \pi]$  such that  $\lambda'(\beta_0) = 0$  or such that  $\lambda(\beta_0)$  is not differentiable in  $\beta = \beta_0$ . These latter points check

$$p_\alpha(\lambda(\beta_0), \beta_0) = 0 \quad \text{and} \quad \frac{\partial p_\alpha}{\partial \lambda}(\lambda(\beta_0), \beta_0) = 0.$$

However, the analysis of these points is extremely complicated and we have restricted ourselves to the study of the value  $\beta_0$  such that  $\lambda'(\beta_0) = 0$ . Therefore, we cannot affirm that our condition is a sufficient condition of stability. Nevertheless, the numerical results presented in section 6 show that this condition is sufficient for numerical experiments. Using the implicit functions theorem, we have

$$\frac{\partial p_\alpha}{\partial \lambda} \lambda'(\beta) + \frac{\partial p_\alpha}{\partial \beta} = 0.$$

Hence  $\beta_0$  satisfies  $\frac{\partial p_\alpha}{\partial \beta}(\lambda(\beta_0), \beta_0) = 0$  and the steps of the proof are, for a given  $\alpha$ :

- (1) Find all the  $\beta_0$  satisfying  $\frac{\partial p_\alpha}{\partial \beta}(\lambda(\beta_0), \beta_0) = 0$ ;
- (2) For each  $\beta_0$ , compute  $\lambda_{\max, \beta_0} = \max \Lambda(\beta_0)$ ;
- (3) Find the maximum value among all the  $\lambda_{\max, \beta_0}$ .

In the following, we only detail the case  $p = 3$ , since the proofs are similar for all  $p = 1, \dots, 5$ .

We have

$$\begin{aligned} \frac{\partial p_\alpha}{\partial \beta}(\lambda, \beta) &= \sin(\beta) \left[ \frac{8}{h^2} (\alpha - 15) \lambda^3 + \frac{240}{h^4} (23 + \alpha - 2 \cos(\beta)) \lambda^2 \right. \\ &\quad \left. + \frac{2880}{h^6} (3\alpha - 65 - 8 \cos(\beta)) \lambda + \frac{100800}{h^8} (-6 \cos(\beta) + 2(\alpha - 3)) \right] \\ &:= \sin(\beta) \tilde{p}_\alpha(\lambda, \beta). \end{aligned}$$

Consequently,

$$\frac{\partial p_\alpha}{\partial \beta}(\beta_0) = 0 \Leftrightarrow \begin{cases} \sin(\beta_0) = 0 \\ \text{or} \\ \tilde{p}_\alpha(\lambda(\beta_0), \beta_0) = 0. \end{cases} \quad (22)$$

First, we consider only the condition  $\sin(\beta_0) = 0$  and we show that it is equivalent to (7). Then, to obtain the necessary and sufficient condition (8), we take into account the condition  $\tilde{p}_\alpha(\lambda(\beta_0), \beta_0) = 0$ .

- If  $\beta_0 = 0$ , we obtain the following eigenvalues:

$$0; \frac{60}{h^2}; \frac{90 + 20\alpha + 2g_1(\alpha)}{h^2}; \frac{90 + 20\alpha - 2g_1(\alpha)}{h^2}$$

where  $g_1(\alpha) = (4545 - 1320\alpha + 100\alpha^2)^{\frac{1}{2}}$ .

It is clear that, for  $\alpha \geq 0$ , the two greatest eigenvalues are  $\lambda_1 = \frac{60}{h^2}$  and  $\lambda_2 = \frac{1}{h^2}(90 + 20\alpha + 2g_1(\alpha))$ . Studying the sign of the quantity

$$h^2(\lambda_1 - \lambda_2) = -150 + 20\alpha + 2g_1(\alpha)$$

we can easily obtain

$$\begin{cases} \lambda_{\max,0} = \lambda_2 & \text{if } \alpha \geq 6 \\ \lambda_{\max,0} = \lambda_1 & \text{if } \alpha < 6. \end{cases} \quad (23)$$

We have proved in section 4.2 that the condition  $\alpha \geq 6$  is a necessary stability condition. Therefore, we only have to consider  $\lambda_{\max,0} = \lambda_2$ .

- If  $\beta_0 = \pi$ , we obtain the following eigenvalues:

$$\frac{2}{h^2}(45 + \sqrt{1605}); \frac{2}{h^2}(45 - \sqrt{1605}); \frac{2}{h^2}(-15 + 6\alpha + g_2(\alpha)); \frac{2}{h^2}(-15 + 6\alpha - g_2(\alpha))$$

where  $g_2(\alpha) = (405 - 240\alpha + 36\alpha^2)^{\frac{1}{2}}$ .

The two greatest eigenvalues are  $\lambda_3 = \frac{2}{h^2}(45 + \sqrt{1605})$  and  $\lambda_4 = \frac{2}{h^2}(-15 + 6\alpha + g_2(\alpha))$ . The study of the sign of  $\lambda_3 - \lambda_4$  implies

$$\begin{cases} \lambda_{\max,\pi} = \lambda_3 & \text{if } \alpha \leq 10 \\ \lambda_{\max,\pi} = \lambda_4 & \text{if } \alpha \geq 10. \end{cases} \quad (24)$$

Now we have to compare  $\lambda_{\max,0}$  and  $\lambda_{\max,\pi}$ . We easily verify that

$$\begin{cases} \lambda_{\max,0} \geq \lambda_{\max,\pi} & \text{if } \alpha \geq \frac{2\sqrt{1605} + 393}{49} \simeq 9.66, \\ \lambda_{\max,\pi} > \lambda_{\max,0} & \text{if } \frac{2\sqrt{1605} + 393}{49} < \alpha. \end{cases} \quad (25)$$

Consequently, considering (23) and (25), a necessary condition of stability is

$$\frac{\Delta t}{h} \leq \begin{cases} \frac{2}{\sqrt{\lambda_{\max,\pi}}} & \text{if } 6 \leq \alpha \leq \alpha_{1,p}, \\ \frac{2}{\sqrt{\lambda_{\max,0}}} & \text{if } \alpha_{1,p} < \alpha, \end{cases} \quad (26)$$

with  $\alpha_{1,p} = \frac{2\sqrt{1605} + 393}{49}$ , which corresponds to the necessary condition (6).

- Let us now find  $\beta_0$  such that  $\tilde{p}_\alpha(\lambda(\beta_0), \beta_0) = 0$ . In fact, we don't have to compute  $\beta_0$ , we interest ourselves only to  $\max \Lambda(\beta_0)$ .

We can easily obtain that

$$\cos(\beta_0) = \frac{(\alpha - 15)h^6\lambda^3(\beta_0) + 30(23 + \alpha)h^4\lambda^2(\beta_0) + 360(3\alpha - 65)h^2\lambda(\beta_0) + 25200(\alpha - 3)}{60(h^4\lambda^2(\beta_0) + 48h^2\lambda(\beta_0) + 1260)} := \tilde{Q}_\alpha(\lambda(\beta_0)).$$

Using this expression of  $\cos(\beta_0)$  in the characteristic polynomial (35) we obtain that  $\lambda = \lambda(\beta_0)$  is solution to

$$q_\alpha(\lambda, \beta_0) = -\frac{1}{15h^8(h^4\lambda^2 + 48h^2\lambda + 1260)} \sum_{i=0}^6 \lambda^i h^{2i} \tilde{c}_i(\alpha) = 0$$

or, equivalently, solution to

$$Q_{p,\alpha}(\lambda) = \sum_{i=0}^6 \lambda^i h^{2i} \tilde{c}_i(\alpha) = 0$$

with

$$\left\{ \begin{array}{l} c_0(\alpha) = 635040000(\alpha^2 - 12\alpha + 36), \\ c_1(\alpha) = 3628800(15\alpha^2 + 70\alpha - 96), \\ c_2(\alpha) = 86400(31\alpha^2 - 447\alpha + 5316), \\ c_3(\alpha) = 14400(8\alpha^2 - 135\alpha - 1728), \\ c_4(\alpha) = 180(17\alpha^2 - 442\alpha + 7740), \\ c_5(\alpha) = 60(\alpha^2 + 16\alpha - 357), \\ c_6(\alpha) = \alpha^2 - 30\alpha + 210. \end{array} \right.$$

After having computed the roots of  $Q_{p,\alpha}$ , we have to verify that  $\beta_0$  is well defined, that is to say  $|\cos(\beta_0)| = |\tilde{Q}_\alpha(\lambda)| \leq 1$ . That is why we are interested uniquely in the eigenvalues  $\lambda$  verifying  $Q_\alpha(\lambda) = 0$  and  $|\tilde{Q}_\alpha(\lambda)| \leq 1$ .  $\square$

## 5. THE $d$ -DIMENSIONAL CASE

In this section, we propose to adapt the technique proposed in [12] to extend the analysis from the 1D case to the  $d$ D case. Here, we only detail the three-dimensional case, since the technique is exactly the same for the two dimensional case.

First of all, we consider an infinite homogeneous 3D domain  $\Omega$  uniformly meshed by cubes of edge  $h$ .

We introduce the following notations.

- $\Omega = \bigcup_{K_J \in \mathcal{T}_h} K_J$  where  $K_J = \prod_{k=1}^3 S_{J_k} = \prod_{k=1}^3 [J_k h, (J_k + 1)h]$  and  $J = (J_k)_{k=1,\dots,3}$ .
- On  $\hat{K} [0, 1]^d$ , we define the Lagrange basis functions  $(\hat{\varphi}_l)_{l \in \{1, \dots, p+1\}^3}$  by

$$\hat{\varphi}_l(\mathbf{x}) = \prod_{k=1}^3 \hat{\varphi}_{l_k}(x_k)$$

where  $\hat{\varphi}_{l_k}$  are the 1D Lagrange basis functions.

- Since the mesh is uniform, the basis functions are defined thanks to the functions  $(\hat{\varphi}_l)_{l \in \{1, \dots, p+1\}^3}$  by

$$\varphi_{\mathbf{m}}^J(\mathbf{x}) = \hat{\varphi}_{\mathbf{m}} \left( \frac{x - Jh}{h} \right) 1_{K_J}(\mathbf{x})$$

where  $1_{K_J}$  is the indicator function of  $K_J$ . These functions can be written as a product of  $d$  1D basis functions:

$$\varphi_{\mathbf{m}}^J(\mathbf{x}) = \prod_{k=1}^3 \hat{\varphi}_{m_k} \left( \frac{x - J_k h}{h} \right) 1_{S_{J_k}}(\mathbf{x}_k).$$

- The different faces of the reference element  $\hat{K}$  are denoted by an exponent  $C$  corresponding to the orientation of the face: North (N), South (S), East (E), West (W), in Front of (F) and in the Back (B) (cf. Fig 1).

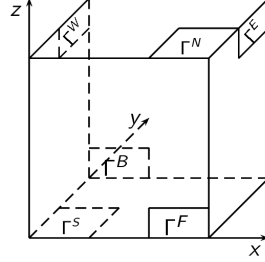


FIGURE 1. Notations of the faces in 3D

Since the mesh is uniform, we can rewrite the problem on an element  $\mathbf{I} = \{I_1, I_2, I_3\}$ :

$$\begin{aligned}
 M_{3,p} \delta^n U_{I_1, I_2, I_3} = & K_{3,p} U_{I_1, I_2, I_3} + K_{3,p}^E U_{I_1+1, I_2, I_3} + K_{3,p}^W U_{I_1-1, I_2, I_3} \\
 & + K_{3,p}^F U_{I_1, I_2+1, I_3} + K_{3,p}^B U_{I_1, I_2-1, I_3} + K_{3,p}^N U_{I_1, I_2, I_3+1} \\
 & + K_{3,p}^S U_{I_1, I_2, I_3-1}
 \end{aligned} \tag{27}$$

where

- $U_{I_1, I_2, I_3}$  corresponds to the restriction of  $U$  on the element  $I$ .
- $\delta^n U_{I_1, I_2, I_3} = \frac{U_{I_1, I_2, I_3}^{n+1} - 2U_{I_1, I_2, I_3}^n + U_{I_1, I_2, I_3}^{n-1}}{\Delta t^2}$
- $M_{3,p}$  is a block of the mass matrix  $M$ ,

$$M_{3,p}(\mathbf{i}, \mathbf{j}) = h^3 \int_{\hat{K}} \hat{\varphi}_{\mathbf{i}} \hat{\varphi}_{\mathbf{j}} d\hat{x}, \quad \mathbf{i}, \mathbf{j} \in \{1, \dots, p+1\}^3$$

- $K_{3,p}$  is a diagonal block of the matrix  $K$

$$\begin{aligned}
 K_{3,p}(\mathbf{i}, \mathbf{j}) = & h \int_{\hat{K}} \nabla \hat{\varphi}_{\mathbf{i}} \cdot \nabla \hat{\varphi}_{\mathbf{j}} d\hat{x} - \sum_{C \in \{N, S, E, W, B, F\}} \frac{h}{2} \int_{\Gamma^C} (\hat{\varphi}_{\mathbf{i}} \nabla \hat{\varphi}_{\mathbf{j}} + \hat{\varphi}_{\mathbf{j}} \nabla \hat{\varphi}_{\mathbf{i}}) \nu_C d\sigma \\
 & + \sum_{C \in \{N, S, E, W, B, F\}} h^2 \int_{\Gamma^C} \gamma \hat{\varphi}_{\mathbf{i}} \hat{\varphi}_{\mathbf{j}} d\sigma, \quad \mathbf{i}, \mathbf{j} \in \{1, \dots, p+1\}^3
 \end{aligned}$$

where  $\nu_C$  is the outward unit normal vector to the face  $\Gamma^C$ .

- $K_{3,p}^C$  is a block of the matrix  $K$  corresponding to the interactions between an element  $I$  and its neighbour on the face  $\Gamma^C$ :

$$\begin{aligned}
K_{3,p}^E(\mathbf{i}, \mathbf{j}) &= \int_{[0,1]^2} \frac{h}{2} (\hat{\varphi}_{\mathbf{i}}(1, x_2, x_3) \nabla \hat{\varphi}_{\mathbf{j}}(0, x_2, x_3) + \hat{\varphi}_{\mathbf{j}}(0, x_2, x_3) \nabla \hat{\varphi}_{\mathbf{i}}(1, x_2, x_3)) \nu_E \\
&\quad - h^2 \hat{\varphi}_{\mathbf{i}}(1, x_2, x_3) \hat{\varphi}_{\mathbf{j}}(0, x_2, x_3) dx_2 dx_3 \\
K_{3,p}^F(\mathbf{i}, \mathbf{j}) &= \int_{[0,1]^2} \frac{h}{2} (\hat{\varphi}_{\mathbf{i}}(x_1, 1, x_3) \nabla \hat{\varphi}_{\mathbf{j}}(x_1, 0, x_3) + \hat{\varphi}_{\mathbf{j}}(x_1, 0, x_3) \nabla \hat{\varphi}_{\mathbf{i}}(x_1, 1, x_3)) \nu_F \\
&\quad - h^2 \hat{\varphi}_{\mathbf{i}}(x_1, 1, x_3) \hat{\varphi}_{\mathbf{j}}(x_1, 0, x_3) dx_1 dx_3 \\
K_{3,p}^N(\mathbf{i}, \mathbf{j}) &= \int_{[0,1]^2} \frac{h}{2} (\hat{\varphi}_{\mathbf{i}}(x_1, x_2, 1) \nabla \hat{\varphi}_{\mathbf{j}}(x_1, x_2, 0) + \hat{\varphi}_{\mathbf{j}}(x_1, x_2, 0) \nabla \hat{\varphi}_{\mathbf{i}}(x_1, x_2, 1)) \nu_N \\
&\quad - h^2 \hat{\varphi}_{\mathbf{i}}(x_1, x_2, 1) \hat{\varphi}_{\mathbf{j}}(x_1, x_2, 0) dx_1 dx_2
\end{aligned}$$

$$\begin{aligned}
K_{3,p}^W((i_1, i_2, i_3), (j_1, j_2, j_3)) &= K_{3,p}^E((j_1, i_2, i_3), (i_1, j_2, j_3)) \\
K_{3,p}^B((i_1, i_2, i_3), (j_1, j_2, j_3)) &= K_{3,p}^F((i_1, j_2, i_3), (j_1, i_2, j_3)) \\
K_{3,p}^S((i_1, i_2, i_3), (j_1, j_2, j_3)) &= K_{3,p}^N((i_1, i_2, j_3), (j_1, j_2, i_3))
\end{aligned}$$

Then, multiplying the equation (27) by the inverse of the mass matrix  $M_{3,p}$ , we obtain

$$\begin{aligned}
\delta^n U_{I_1, I_2, I_3} &= N_{3,p} U_{I_1, I_2, I_3} + N_{3,p}^E U_{I_1+1, I_2, I_3} + N_{3,p}^W U_{I_1-1, I_2, I_3} \\
&\quad + N_{3,p}^F U_{I_1, I_2+1, I_3} + N_{3,p}^B U_{I_1, I_2-1, I_3} + N_{3,p}^N U_{I_1, I_2, I_3+1} \\
&\quad + N_{3,p}^S U_{I_1, I_2, I_3-1}
\end{aligned} \tag{28}$$

where  $N_{3,p} = M_{3,p}^{-1} K_{3,p}$  and  $N_{3,p}^C = M_{3,p}^{-1} K_{3,p}^C$ .

Now, we are interested in rewriting the matrices  $N_{3,p}$  and  $N_{3,p}^C$  with respect to the matrices we have obtained for the one dimensional case.

### 5.1. From the 3 dimensional case to the one dimensional case

The coefficients of  $M_{3,p}$ ,  $K_{3,p}$ ,  $K_{3,p}^C$ ,  $N_{3,p}$  and  $N_{3,p}^C$  can be deduced from the coefficients of  $M_{1,p}$ ,  $K_{1,p}$ ,  $K_{1,p}^W$ ,  $N_{1,p}$  and  $N_{1,p}^W$  thanks to the following theorem.



**Theorem 5.1.** For all  $\mathbf{m} = (m_k)_{k=1,\dots,3} \in \{1, \dots, p+1\}^3$  and  $\mathbf{n} = (n_k)_{k=1,\dots,3} \in \{1, \dots, p+1\}^3$ , we have

$$\begin{aligned}
1. \quad & M_{3,p}(\mathbf{m}, \mathbf{n}) = \prod_{i=1}^3 M_{1,p}(m_i, n_i), \\
2. \quad & K_{3,p}(\mathbf{m}, \mathbf{n}) = \sum_{i=1}^3 \left( K_{1,p}(m_i, n_i) \prod_{k=1, k \neq i}^3 M_{1,p}(m_k, n_k) \right), \\
3. \quad & K_{3,p}^C(\mathbf{m}, \mathbf{n}) = K_{1,p}^W(m_{p_C}, n_{p_C}) \prod_{k=1, k \neq p_C}^3 M_{1,p}(m_k, n_k), \\
4. \quad & N_{3,p}(\mathbf{m}, \mathbf{n}) = \sum_{p=1}^3 \left( N_{1,p}(m_p, n_p) \prod_{k=1, k \neq p}^3 \delta_{m_k, n_k} \right), \\
5. \quad & N_{3,p}^C(\mathbf{m}, \mathbf{n}) = N_{1,p}^W(m_{p_C}, n_{p_C}) \prod_{k=1, k \neq p_C}^3 \delta_{m_k, n_k},
\end{aligned} \tag{29}$$

$$\text{where } p_C = \begin{cases} 1 & \text{if } C \in \{E, W\}, \\ 2 & \text{if } C \in \{N, S\}, \\ 3 & \text{if } C \in \{B, F\}, \end{cases} \quad \text{and } N_{1,p} = M_{1,p}^{-1} K_{1,p} \text{ and } N_{1,p}^C = M_{1,p}^{-1} K_{1,p}^C.$$

The proof of this theorem is given in appendix C.

## 5.2. Consequences on the stability analysis

Let us now apply a Fourier transform in the three directions to (28) to obtain, for  $\beta = [-\pi, \pi]^3$ ,

$$\delta^n \tilde{U}_{\beta_1, \beta_2, \beta_3} = \mathbf{N}_\beta \tilde{U}_{\beta_1, \beta_2, \beta_3} \tag{30}$$

where the matrix  $\mathbf{N}_\beta$  is defined by

$$\mathbf{N}_\beta(\mathbf{m}, \mathbf{n}) = \sum_{j=1}^3 \left[ N_{1,p}(m_j, n_j) \prod_{q=1, q \neq j}^3 \delta_{m_q, n_q} + e^{i\beta_j} N_{1,p}^W(m_j, n_j) \prod_{q=1, q \neq j}^3 \delta_{m_q, n_q} + e^{-i\beta_j} N_{1,p}^W(n_j, m_j) \prod_{q=1, q \neq j}^3 \delta_{m_q, n_q} \right]$$

which can be rewritten as

$$\begin{aligned}
\mathbf{N}_\beta(\mathbf{m}, \mathbf{n}) &= \sum_{j=1}^3 \left( (N_{1,p}(m_j, n_j) + e^{i\beta_j} N_{1,p}^W(m_j, n_j) + e^{-i\beta_j} N_{1,p}^W(n_j, m_j)) \prod_{q=1, q \neq j}^3 \delta_{m_q, n_q} \right) \\
&= \sum_{j=1}^3 N_{\beta_j}(m_j, n_j) \prod_{q=1, q \neq j}^3 \delta_{m_q, n_q}.
\end{aligned}$$

Using the stability analysis as in section 4.1, the stability of the scheme is ensured if and only if

$$\lambda_{\min,3} \geq 0 \quad \text{and} \quad \lambda_{\max,3} \leq \frac{4}{\Delta t^2}$$

where  $\lambda_{\min,3} = \min_{\beta \in [-\pi, \pi]^3} (\min \Lambda(\mathbf{N}_\beta))$ ,  $\lambda_{\max,3} = \max_{\beta \in [-\pi, \pi]^3} (\max \Lambda(\mathbf{N}_\beta))$  and  $\Lambda(\mathbf{N}_\beta)$  is the set of eigenvalues of  $\mathbf{N}_\beta$ .

To compute these values, we use the following lemma.

**Lemma 5.2.** *Let  $(\lambda_i^\beta)_{i=1,\dots,p+1}$  denote the eigenvalues of  $N_\beta$  and  $(v_i^\beta)_{i=1,\dots,p+1}$  the associated eigenvectors. Then, the eigenvalues of  $\mathbf{N}_\beta$  are given by*

$$\lambda_{\mathbf{i}}^\beta = \sum_{k=1}^3 \lambda_{i_k}^{\beta_k}$$

and the associated eigenvectors by

$$v_{\mathbf{i}}^\beta(\mathbf{m}) = \prod_{k=1}^3 v_{i_k}^{\beta_k}(m_k). \quad (31)$$

*Proof.* Let  $v_{\mathbf{i}}^\beta$  defined by (31), then

$$\begin{aligned} (N_\beta v_{\mathbf{i}}^\beta)(\mathbf{m}) &= \sum_{n_1, \dots, n_3=1}^{p+1} \left( \sum_{q=1}^3 N_{\beta_q}(m_q, n_q) \prod_{k=1, k \neq q}^3 \delta_{m_k, n_k} \prod_{k=1}^3 v_{i_k}^{\beta_k}(n_k) \right) \\ &= \sum_{q=1}^3 \left( \left( \sum_{n_q=1}^{p+1} N_{\beta_q}(m_q, n_q) v_{i_q}^{\beta_q}(n_q) \right) \prod_{k=1, k \neq q}^3 v_{i_k}^{\beta_k}(m_k) \right) \\ &= \sum_{q=1}^3 \left( \lambda_{i_q}^{\beta_q} v_{i_q}^{\beta_q}(m_q) \prod_{k=1, k \neq q}^3 v_{i_k}^{\beta_k}(m_k) \right) \\ &= \left( \sum_{q=1}^3 \lambda_{i_q}^{\beta_q} \right) v_{\mathbf{i}}^\beta(\mathbf{m}) \end{aligned}$$

□

It is then clear that

$$\lambda_{\min,3} = 3\lambda_{\min} \quad \text{and} \quad \lambda_{\max,3} = 3\lambda_{\max}.$$

Hence, the scheme (31) is stable, if and only if

$$\lambda_{\min} \geq 0 \quad \text{and} \quad \frac{c\Delta t}{h} \leq \frac{1}{\sqrt{3}} \sqrt{\frac{2}{\lambda_{\max}}}.$$

The first condition is equivalent to condition (6), while the second one implies (7) and (8).

### 5.3. Extension to rectangular or parallelepiped mesh

For the sake of simplicity, we restricted our theorem to the case of squared or cubic mesh. However, one can extend the proof to the case of rectangular or parallelepipeds meshes to show that a necessary stability condition is in 2D:

$$\gamma \geq \frac{p(p+1)}{2 \min(h_x, h_y)}$$

and in 3D:

$$\gamma \geq \frac{p(p+1)}{2 \min(h_x, h_y, h_z)}.$$

Here,  $h_x$ ,  $h_y$  and  $h_z$  denote respectively the length of the edges of the elements in the  $x$ ,  $y$  and  $z$  direction. The minimal value of  $h_x$ ,  $h_y$  and  $h_z$  is actually the diameter of the inscribed sphere of each element. This remark confirms that the third definition of  $\xi_F$  using the diameter of the inscribed sphere or circle in 2D is the most appropriate.

The proof could also be extended to obtain a CFL condition, but its expression is complicated and does not add much insight.

## 6. NUMERICAL RESULTS

In this section, we first represent the behaviour of the CFL condition with respect to  $\alpha$  and we show that the set  $V_{p,\alpha}$  is empty for almost all the values of  $\alpha$  (section 6.1). This illustrates the fact that theorem 3.1 is actually necessary and sufficient for most of the values of  $\alpha$ . Then, we compare our analytical CFL condition in infinite domain to the CFL condition computed numerically on finite meshes in order to illustrate the validity of theorem 3.3 (section 6.2).

### 6.1. Behaviour of the CFL condition with respect to $\alpha$

In Fig. 2, 4, 6, 8 and 10 we plot the functions  $C_{1,p}$  (blue line with diamonds),  $C_{2,p}(\alpha)$  (red line with circles) and  $C_{3,p}(\alpha)$  (black line) respectively for  $p = 1, 2, 3, 4$  and 5. The function  $C_{3,p}(\alpha)$  only modifies the CFL condition in a small segment around  $\alpha_{1,p}$ . The behaviour is confirmed in Fig. 3, 5, 7, 9 and 11 which represent a zoom around  $\alpha_{1,p}$ . These numerical results confirm the fact that theorem 3.1 provides actually a necessary and sufficient condition except for a small range of  $\alpha$ . Moreover, the CFL condition remains constant for  $\alpha$  from  $\frac{p(p+1)}{2}$  to a close value of  $\alpha_p^1$ , which means that it is not necessary to choose  $\alpha = \frac{p(p+1)}{2}$  to optimize the time step.

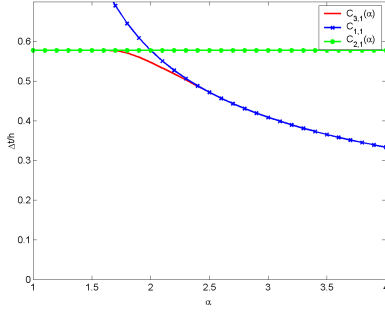


FIGURE 2. The 3 conditions for  $p = 1$

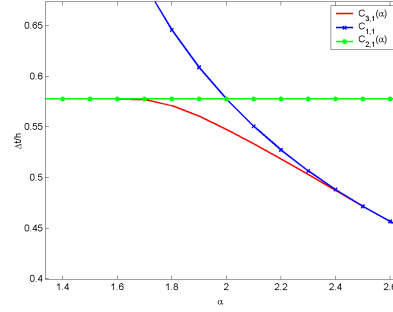


FIGURE 3. Zoom on the 3 cond. for  $p = 1$

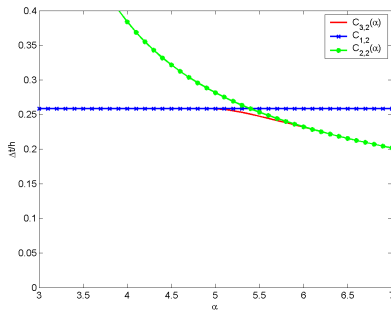


FIGURE 4. The 3 conditions for  $p = 2$

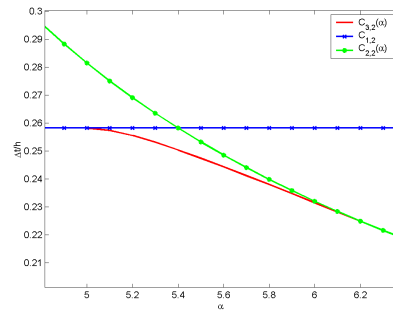


FIGURE 5. Zoom on the 3 cond. for  $p = 2$

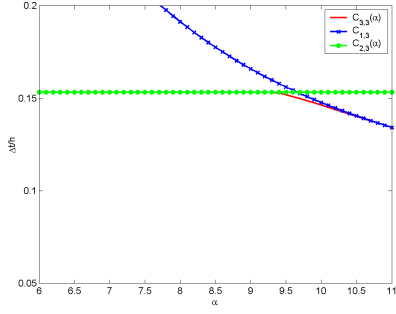


FIGURE 6. The 3 conditions for  $p = 3$

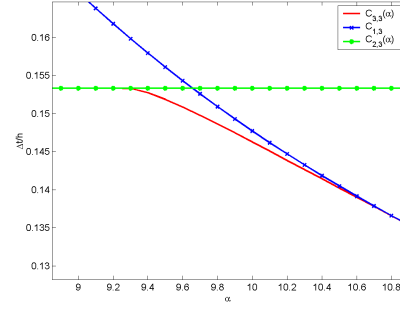


FIGURE 7. Zoom on the 3 cond. for  $p = 3$

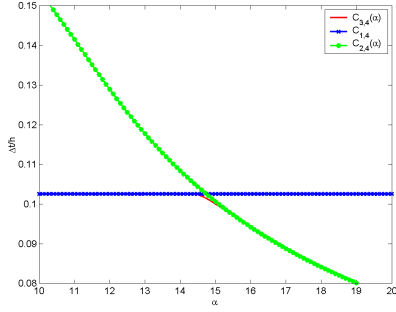


FIGURE 8. The 3 conditions for  $p = 4$

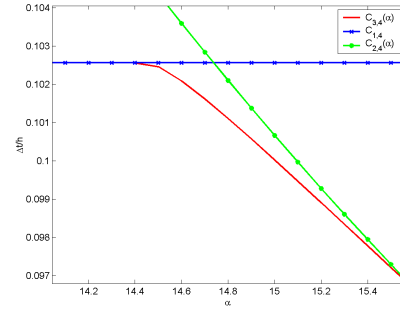


FIGURE 9. Zoom on the 3 cond. for  $p = 4$

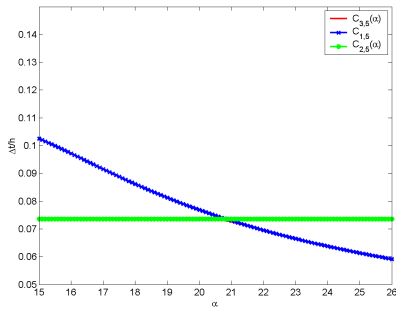


FIGURE 10. The 3 conditions for  $p = 5$

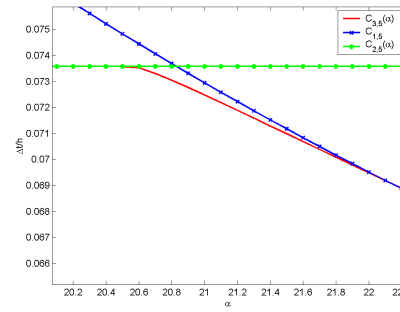


FIGURE 11. Zoom on the 3 cond. for  $p = 5$

## 6.2. Comparison with numerical experiments

In this section, we compare the results we have obtained previously with numerical experiments. We consider the simulation of wave propagation in an homogeneous 1D domain  $\Omega = [0, 10]$  with a velocity  $c = (\mu/\rho)^{1/2} = 1 \text{ ms}^{-1}$ . We impose also Dirichlet boundary conditions at the both ends of the domain and the length of the space step is  $h = 0.1$ .

We computed numerically the greatest eigenvalue  $\lambda_{\max}$  of the matrix  $M^{-1}K$  and we deduced the CFL condition of the scheme using the formula  $\frac{c\Delta t}{h} \leq \frac{2}{\sqrt{\lambda_{\max}}}$ . In Fig. 12, 13, 14, 15 and 16 we compare the analytical CFL (red line) obtained by theorem 3.3 to the numerical CFL (triangles), respectively for  $p = 1, 2, 3, 4$  and 5. All figures show a very good agreement between the analytical and the numerical CFL.

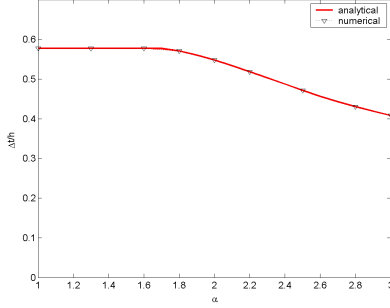


FIGURE 12. Numerical comparison in  $P^1$

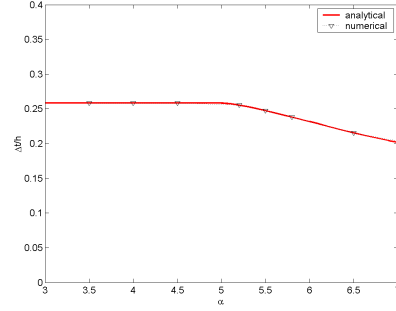


FIGURE 13. Numerical comparison in  $P^2$

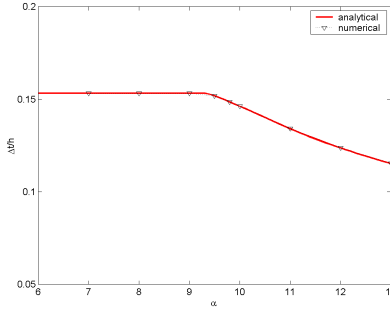


FIGURE 14. Numerical comparison in  $P^3$

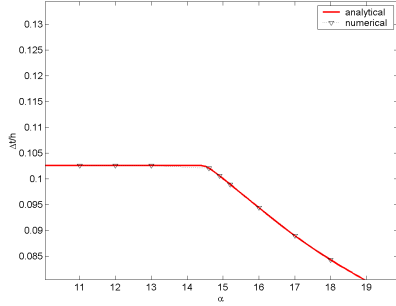
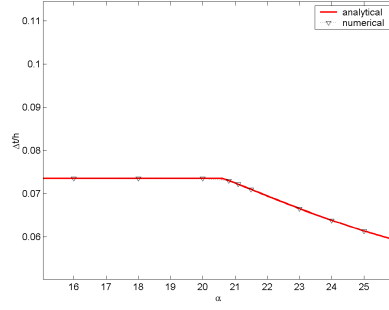


FIGURE 15. Numerical comparison in  $P^4$

## 7. CONCLUSION

In this paper, we have proposed necessary conditions of  $L^2$ -stability of an IPDG method using regular meshes and the numerical results showed that these conditions are actually sufficient in practice. It also confirm the conjecture of Ainsworth, Monk and Muniz up to  $p = 5$ . Moreover, we have observed that the CFL condition is constant with respect to  $\alpha$  on a segment  $\left[\frac{p(p+1)}{2}, \tilde{\alpha}\right]$  and is decreasing as  $\alpha^{-1/2}$  for  $\alpha > \tilde{\alpha}$ . This means that it is not necessary to choose  $\alpha$  too close to  $\frac{p(p+1)}{2}$  to improve the CFL condition. Finally, we have observed that a good choice for  $\xi_F$  should be the diameter of the inscribed circle (or sphere). This should be confirmed by an analysis on triangular meshes, which will be the topic of a future work.

FIGURE 16. Numerical comparison in  $P^5$ APPENDIX A. EXPRESSION OF THE POLYNOMIAL  $p_\alpha$ 

- In the case of discontinuous basis functions of degree 1, we can easily obtain the following characteristic polynomial associated to the matrix  $A_\beta$

$$p_\alpha(\lambda, \beta) = \lambda^2 + c_1(\alpha, \beta)\lambda + c_0(\alpha, \beta) \quad (32)$$

with

$$\begin{cases} c_1(\alpha, \beta) = \frac{4}{h^2} ((3 - \alpha) \cos(\beta) - 2\alpha) \\ c_0(\alpha, \beta) = \frac{12}{h^4} (\cos^2(\beta) - 2\alpha \cos(\beta) + 2\alpha - 1). \end{cases}$$

- In the case of discontinuous basis functions of degree 2, the characteristic polynomial associated to the matrix  $A_\beta$  is

$$p_\alpha(\lambda, \beta) = -\lambda^3 + c_2(\alpha, \beta)\lambda^2 + c_1(\alpha, \beta)\lambda + c_0(\alpha, \beta) \quad (33)$$

with

$$\begin{cases} c_2(\alpha, \beta) = -\frac{6}{h^2} ((\alpha - 8) \cos(\beta) - 3\alpha) \\ c_1(\alpha, \beta) = \frac{12}{h^4} (-6 \cos^2(\beta) - 2(15 + 4\alpha) \cos(\beta) + 4(24 - 13\alpha)) \\ c_0(\alpha, \beta) = -\frac{1440}{h^6} (\cos^2(\beta) + (\alpha - 3) \cos(\beta) + 2 - \alpha). \end{cases}$$

- In the case of discontinuous basis functions of degree 4, the characteristic polynomial associated to the matrix  $A_\beta$  is

$$p_\alpha(\lambda, \beta) = -\lambda^5 + c_4(\alpha, \beta)\lambda^4 + c_3(\alpha, \beta)\lambda^3 + c_2(\alpha, \beta)\lambda^2 + c_1(\alpha, \beta)\lambda + c_0(\alpha, \beta) \quad (34)$$

with

$$\left\{ \begin{array}{l} c_4(\alpha, \beta) = -\frac{10}{h^2} ((\alpha - 24) \cos(\beta) - 5\alpha) \\ c_3(\alpha, \beta) = -\frac{120}{h^4} (5 \cos^2(\beta) + (4\alpha + 287) \cos(\beta) + 10(15\alpha - 88)\alpha) \\ c_2(\alpha, \beta) = -\frac{10080}{h^6} (5 \cos^2(\beta) + (3\alpha - 305) \cos(\beta) + 990 - 133\alpha) \\ c_1(\alpha, \beta) = -\frac{201600}{h^8} (15 \cos^2(\beta) + (8\alpha + 165) \cos(\beta) + 2(59\alpha - 468)) \\ c_0(\alpha, \beta) = -\frac{50803200}{h^{10}} (2 \cos^2(\beta) + (\alpha - 10) \cos(\beta) + 8 - \alpha) . \end{array} \right.$$

- In the case of discontinuous basis functions of degree 5, the characteristic polynomial associated to the matrix  $A_\beta$  is

$$p_\alpha(\lambda, \beta) = \lambda^6 + c_5(\alpha, \beta) \lambda^5 + c_4(\alpha, \beta) \lambda^4 + c_3(\alpha, \beta) \lambda^3 + c_2(\alpha, \beta) \lambda^2 + c_1(\alpha, \beta) \lambda + c_0(\alpha, \beta) \quad (35)$$

with

$$\left\{ \begin{array}{l} c_5(\alpha, \beta) = -\frac{12}{h^2} ((\alpha - 35) \cos(\beta) + 6\alpha) \\ c_4(\alpha, \beta) = -\frac{420}{h^4} (-3 \cos^2(\beta) + (2\alpha + 336) \cos(\beta) + 1155 - 134\alpha) \\ c_3(\alpha, \beta) = -\frac{40320}{h^6} (-4 \cos^2(\beta) + (2\alpha - 702) \cos(\beta) + 256\alpha - 2849\alpha) \\ c_2(\alpha, \beta) = -\frac{1814400}{h^8} (-9 \cos^2(\beta) + (4\alpha + 770) \cos(\beta) + 9343 - 326\alpha) \\ c_1(\alpha, \beta) = -\frac{101606400}{h^{10}} (-12 \cos^2(\beta) + (5\alpha - 303) \cos(\beta) + 94\alpha - 1170) \\ c_0(\alpha, \beta) = -\frac{10059033600}{h^{12}} (-5 \cos^2(\beta) + (2\alpha - 20) \cos(\beta) + 25 - 2\alpha) . \end{array} \right.$$

## APPENDIX B. DEFINITION OF $Q_{p,\alpha}$ AND $\tilde{Q}_{p,\alpha}$

We present here the expressions of the polynomial  $Q_{p,\alpha}$  and the rational function  $\tilde{Q}_{p,\alpha}$  for  $1 \leq p \leq 5$ .

- For the polynomials of degree 1,  $\tilde{Q}_{p,\alpha}$  is defined by

$$\tilde{Q}_{p,\alpha}(\lambda) = \frac{h^2 \lambda}{2} \left( \frac{\alpha}{3} - 1 \right) + \alpha.$$

We have,

$$Q_{p,\alpha}(\lambda) = \sum_{i=0}^2 \lambda^i h^{2i} \tilde{c}_i(\alpha)$$

with

$$\left\{ \begin{array}{l} \tilde{c}_0(\alpha) = 36(\alpha^2 - 2\alpha + 1), \\ \tilde{c}_1(\alpha) = 12(\alpha^2 - \alpha), \\ \tilde{c}_2(\alpha) = \alpha^2 - 6\alpha + 6, \end{array} \right.$$

- In the case  $p = 2$ , the definition of  $\tilde{Q}_{p,\alpha}$  is

$$\tilde{Q}_{p,\alpha}(\lambda) = -\frac{(\alpha - 1)h^4\lambda^2 + 4(15 + 4\alpha)h^2\lambda + 240(\alpha - 3)}{24(h^2\lambda + 20)}.$$

The polynomial  $Q_{p,\alpha}$  is such that

$$Q_{p,\alpha}(\lambda) = \sum_{i=0}^4 \lambda^i h^{2i} \tilde{c}_i(\alpha)$$

with

$$\begin{cases} \tilde{c}_0(\alpha) = 57600(\alpha^2 - 2\alpha + 1), \\ \tilde{c}_1(\alpha) = 1920(4\alpha^2 - 43\alpha + 39), \\ \tilde{c}_2(\alpha) = 16(46\alpha^2 - 342\alpha + 1521), \\ \tilde{c}_3(\alpha) = 8(4\alpha^2 + \alpha - 140), \\ \tilde{c}_4(\alpha) = \alpha^2 - 16\alpha + 56. \end{cases}$$

- For  $p = 3$ , we have

$$Q_{p,\alpha}(\lambda) = \sum_{i=0}^6 \lambda^i h^{2i} \tilde{c}_i(\alpha)$$

with

$$\begin{cases} \tilde{c}_0(\alpha) = 635040000(\alpha^2 - 12\alpha + 36), \\ \tilde{c}_1(\alpha) = 3628800(15\alpha^2 + 70\alpha - 96), \\ \tilde{c}_2(\alpha) = 86400(31\alpha^2 - 447\alpha + 5316), \\ \tilde{c}_3(\alpha) = 14400(8\alpha^2 - 135\alpha - 1728), \\ \tilde{c}_4(\alpha) = 180(17\alpha^2 - 442\alpha + 7740), \\ \tilde{c}_5(\alpha) = 60(\alpha^2 + 16\alpha - 357), \\ \tilde{c}_6(\alpha) = \alpha^2 - 30\alpha + 210. \end{cases}$$

and  $\tilde{Q}_{p,\alpha}$  is defined by

$$\tilde{Q}_{p,\alpha}(\lambda) = \frac{(\alpha - 15)h^6\lambda^3 + 30(23 + \alpha)h^4\lambda^2 + 360(3\alpha - 65)h^2\lambda + 25200(\alpha - 3)}{60(h^4\lambda^2 + 48h^2\lambda + 1260)}.$$

- For the polynomials of degree 4,  $\tilde{Q}_{p,\alpha}$  is defined by

$$\tilde{Q}_{p,\alpha}(\lambda) = -\frac{\tilde{A}_{p,\alpha}(\lambda)}{120(169344 + h^6\lambda^3 + 84h^4\lambda^2 + 5040h^2\lambda)}$$

with

$$\begin{aligned} \tilde{A}_{p,\alpha}(\lambda) = & \lambda^4 h^8 (\alpha - 24) + 12\lambda^3 h^6 (4\alpha + 287) + 1008\lambda^2 h^4 (3\alpha - 305) \\ & + 20160\lambda h^2 (8\alpha - 165) + 5080320(\alpha - 1) \end{aligned}$$



And, we have

$$Q_{p,\alpha}(\lambda) = \sum_{i=0}^8 \lambda^i h^{2i} \tilde{c}_i(\alpha).$$

with

$$\left\{ \begin{array}{l} \tilde{c}_0(\alpha) = 25809651302400(\alpha - 6)^2, \\ \tilde{c}_1(\alpha) = 204838502400(8\alpha^2 - 357\alpha + 1854), \\ \tilde{c}_2(\alpha) = 81285120(698\alpha^2 + 3882\alpha + 292185), \\ \tilde{c}_3(\alpha) = 203212800(72\alpha^2 - 13791\alpha - 328), \\ \tilde{c}_4(\alpha) = 48384(719\alpha^2 - 12750\alpha + 2419275), \\ \tilde{c}_5(\alpha) = 8064(76\alpha^2 - 972\alpha - 286209), \\ \tilde{c}_6(\alpha) = 144(58\alpha^2 - 5282\alpha + 201609), \\ \tilde{c}_7(\alpha) = 24(4\alpha^2 + 241\alpha - 6972), \\ \tilde{c}_8(\alpha) = \alpha^2 - 48\alpha + 55. \end{array} \right.$$

- For  $p = 5$ ,  $\tilde{Q}_{p,\alpha}$  is such that

$$\tilde{Q}_{p,\alpha}(\lambda) = \frac{\tilde{A}_{p,\alpha}(\lambda)}{210(39916800 + \lambda^4 h^8 + 128\lambda^3 h^6 + 12960\lambda^2 h^4 + 967680\lambda h^2)}$$

with

$$\begin{aligned} \tilde{A}_{p,\alpha}(\lambda) = & \lambda^5 h^{10}(\alpha - 35) + 70\lambda^4 h^8(\alpha + 168) + 6720\lambda^3 h^6(\alpha - 351) \\ & + 302400\lambda^2 h^4(2\alpha + 385) + 8467200\lambda h^2(5\alpha - 303) + 1676505600(\alpha - 1) \end{aligned}$$

and we have

$$Q_{p,\alpha}(\lambda) = \sum_{i=0}^{10} \lambda^i h^{2i} \tilde{c}_i(\alpha)$$

with

$$\left\{ \begin{array}{l} \tilde{c}_0(\alpha) = 2810671026831360000(\alpha - 15)^2, \\ \tilde{c}_1(\alpha) = 28390616432640000(5\alpha + 168)(\alpha - 15), \\ \tilde{c}_2(\alpha) = 10241925120000(373\alpha^2 - 35067\alpha + 855423), \\ \tilde{c}_3(\alpha) = 3072577536000(24\alpha^2 - 895\alpha - 159240), \\ \tilde{c}_4(\alpha) = 1016064000(1151\alpha^2 - 11360\alpha + 24379995), \\ \tilde{c}_5(\alpha) = 67737600(257\alpha^2 - 3570\alpha - 9096540), \\ \tilde{c}_6(\alpha) = 2822400(76\alpha^2 - 2417\alpha + 2988895), \\ \tilde{c}_7(\alpha) = 67200(32\alpha^2 + 2383\alpha - 974820), \\ \tilde{c}_8(\alpha) = 140(131\alpha^2 - 37062\alpha + 2285235), \\ \tilde{c}_9(\alpha) = 140(\alpha^2 + 151\alpha - 5912), \\ \tilde{c}_{10}(\alpha) = \alpha^2 - 70\alpha + 1190. \end{array} \right.$$

### APPENDIX C. PROOF OF THEOREM 5.1

This section is devoted to the proof of the following theorem

**Theorem C.1.** *For all  $\mathbf{m} = (m_k)_{k=1,\dots,3} \in \{1, \dots, p+1\}^3$  and  $\mathbf{n} = (n_k)_{k=1,\dots,3} \in \{1, \dots, p+1\}^3$ , we have*

$$\begin{aligned} 1. \quad & M_{3,p}(\mathbf{m}, \mathbf{n}) = \prod_{i=1}^3 M_{1,p}(m_i, n_i), \\ 2. \quad & K_{3,p}(\mathbf{m}, \mathbf{n}) = \sum_{i=1}^3 \left( K_{1,p}(m_i, n_i) \prod_{k=1, k \neq i}^3 M_{1,p}(m_k, n_k) \right) \\ 3. \quad & K_{3,p}^C(\mathbf{m}, \mathbf{n}) = K_{1,p}^W(m_{p_C}, n_{p_C}) \prod_{k=1, k \neq p_C}^3 M_{1,p}(m_k, n_k) \\ 4. \quad & N_{3,p}(\mathbf{m}, \mathbf{n}) = \sum_{p=1}^3 N_{1,p}(m_p, n_p) \prod_{k=1, k \neq p}^3 \delta_{m_k, n_k} \\ 5. \quad & N_{3,p}^C(\mathbf{m}, \mathbf{n}) = N_{1,p}^W(m_{p_C}, n_{p_C}) \prod_{k=1, k \neq p_C}^3 \delta_{m_k, n_k} \end{aligned} \tag{36}$$

$$\text{where } p_C = \begin{cases} 1 & \text{if } C \in \{E, W\}, \\ 2 & \text{if } C \in \{N, S\}, \\ 3 & \text{if } C \in \{B, F\}, \end{cases} \quad \text{and } N_{1,p} = M_{1,p}^{-1} K_{1,p} \text{ and } N_{1,p}^C = M_{1,p}^{-1} K_{1,p}^C.$$

*Proof.*

- *Proof of 1.*

Considering the notations and the results of the section 5, we have

$$\begin{aligned}
 M_{3,p}(\mathbf{m}, \mathbf{n}) &= h^3 \int_{\hat{K}} \hat{\varphi}_{\mathbf{m}} \hat{\varphi}_{\mathbf{n}} d\mathbf{x} \\
 &= h^3 \int_{\hat{K}} \prod_{i=1}^3 \hat{\varphi}_{m_i}(x_i) \prod_{i=1}^3 \hat{\varphi}_{n_i}(x_i) \prod_{i=1}^3 dx_i \\
 &= \prod_{i=1}^3 h \int_{[0,1]} \hat{\varphi}_{m_i}(x_i) \hat{\varphi}_{n_i}(x_i) dx_i \\
 &= \prod_{i=1}^3 M_{1,p}(m_i, n_i)
 \end{aligned}$$

- *Proof of 2.*

We first have the following lemma for the volumic term.

**Lemma C.2.** *For all  $\mathbf{m} = (m_k)_{k=1,\dots,3} \in \{1, \dots, p+1\}^3$  and  $\mathbf{n} = (n_k)_{k=1,\dots,3} \in \{1, \dots, p+1\}^3$ , we have*

$$h \int_{\hat{K}} \nabla \hat{\varphi}_{\mathbf{m}} \cdot \nabla \hat{\varphi}_{\mathbf{n}} d\mathbf{x} = \sum_{i=1}^3 \left( \frac{1}{h} \int_{[0,1]} \frac{\partial \hat{\varphi}_{m_i}}{\partial x_i}(x_i) \frac{\partial \hat{\varphi}_{n_i}}{\partial x_i}(x_i) dx_i \prod_{k=1, k \neq i}^3 M_{1,p}(m_k, n_k) \right).$$

*Proof.* We know that,  $\forall \mathbf{m} \in \{1, \dots, p+1\}^d$

$$\hat{\varphi}_{\mathbf{m}}(\mathbf{x}) = \prod_{k=1}^3 \hat{\varphi}_{m_k}(x_k)$$

which implies that

$$\frac{\partial \hat{\varphi}_{\mathbf{m}}}{\partial x_k}(\mathbf{x}) = \frac{\partial \hat{\varphi}_{m_k}}{\partial x_k}(x_k) \prod_{i=1, i \neq k}^3 \hat{\varphi}_{m_i}(x_i).$$

Then, using the same reasoning as previously,

$$\begin{aligned}
 h \int_{\hat{K}} \nabla \hat{\varphi}_{\mathbf{m}} \cdot \nabla \hat{\varphi}_{\mathbf{n}} d\mathbf{x} &= \sum_{i=1}^3 \left( \frac{1}{h} \int_{[0,1]} \frac{\partial \hat{\varphi}_{m_i}}{\partial x_i}(x_i) \frac{\partial \hat{\varphi}_{n_i}}{\partial x_i}(x_i) dx_i \prod_{k=1, k \neq i}^3 h \int_{[0,1]} \hat{\varphi}_{m_k}(x_k) \hat{\varphi}_{n_k}(x_k) dx_k \right) \\
 &= \sum_{i=1}^3 \left( \frac{1}{h} \int_{[0,1]} \frac{\partial \hat{\varphi}_{m_i}}{\partial x_i}(x_i) \frac{\partial \hat{\varphi}_{n_i}}{\partial x_i}(x_i) dx_i \prod_{k=1, k \neq i}^3 M_{1,p}(m_k, n_k) \right)
 \end{aligned}$$

which ends the proof.  $\square$

Now, we have to deal with the surface terms.

Let us first remark that, over all the faces  $\Gamma^C$ ,

$$\nabla \hat{\varphi}_{\mathbf{m}} \cdot \nu|_{\Gamma^C} = \frac{\partial \hat{\varphi}_{m_{p_C}}}{\partial x_{p_C}}(x_{p_C}) \nu_{1,C} \prod_{k=1, k \neq p_C}^3 \hat{\varphi}_{m_k}(x_k) \quad (37)$$

where  $\nu_{1,C}$  is the outward unit normal vector in the one dimensional case defined by

$$\nu_{1,C} = \begin{cases} 1 & \text{if } C \in \{E, N, F\}, \\ -1 & \text{if } C \in \{W, S, B\} \end{cases}$$

and  $x_{p_C}$  is defined by

$$x_{p_C} = \begin{cases} 1 & \text{if } C \in \{E, N, F\}, \\ 0 & \text{if } C \in \{W, S, B\}. \end{cases}$$

Then, we can propose the following lemma.

**Lemma C.3.** *For all  $C \in \{E, W, N, S, F, B\}$ , we have  $\forall \mathbf{m}, \mathbf{n} \in \{1, \dots, p+1\}^d$*

$$\begin{aligned} h \int_{\Gamma^C} \hat{\varphi}_{\mathbf{m}} (\nabla \hat{\varphi}_{\mathbf{n}} \cdot \nu) d\sigma &= \frac{1}{h} \hat{\varphi}_{m_{p_C}} (x_{p_C}) \frac{\partial \hat{\varphi}_{n_{p_C}}}{\partial x_{p_C}} (x_{p_C}) \nu_{1,C} \prod_{k=1, k \neq p_C}^3 M_{1,p} (m_k, n_k) \\ h^2 \int_{\Gamma^C} \gamma \hat{\varphi}_{\mathbf{m}} \hat{\varphi}_{\mathbf{n}} d\sigma &= \gamma \hat{\varphi}_{m_{p_C}} (x_{p_C}) \hat{\varphi}_{n_{p_C}} (x_{p_C}) \prod_{k=1, k \neq p_C}^3 M_{1,p} (m_k, n_k). \end{aligned}$$

*Proof.* First of all, using (37), we have

$$\begin{aligned} h \int_{\Gamma^C} \hat{\varphi}_{\mathbf{m}} (\nabla \hat{\varphi}_{\mathbf{n}} \cdot \nu) d\sigma &= h \int_{\Gamma^C} \left( \prod_{k=1}^3 \hat{\varphi}_{m_k} (x_k) \right) \left( \frac{\partial \hat{\varphi}_{n_{p_C}}}{\partial x_{p_C}} (x_{p_C}) \nu_{1,C} \prod_{k=1, k \neq p_C}^3 \hat{\varphi}_{n_k} (x_k) \right) \prod_{k=1, k \neq p_C}^3 dx_k \\ &= \frac{1}{h} \hat{\varphi}_{m_{p_C}} (x_{p_C}) \frac{\partial \hat{\varphi}_{n_{p_C}}}{\partial x_{p_C}} (x_{p_C}) \nu_{1,C} \prod_{k=1, k \neq p_C}^3 h^2 \int_{[0,1]} \hat{\varphi}_{m_k} (x_k) \hat{\varphi}_{n_k} (x_k) dx_k \end{aligned}$$

which can be rewritten as

$$\int_{\Gamma^C} \hat{\varphi}_{\mathbf{m}} (\nabla \hat{\varphi}_{\mathbf{n}} \cdot \nu) d\sigma = \hat{\varphi}_{m_{p_C}} (x_{p_C}) \frac{\partial \hat{\varphi}_{n_{p_C}}}{\partial x_{p_C}} (x_{p_C}) \nu_{1,C} \prod_{k=1, k \neq p_C}^3 M_{1,p} (m_k, n_k).$$

In the same way, for the penalization term, we have:

$$\begin{aligned} h^2 \int_{\Gamma^C} \gamma \hat{\varphi}_{\mathbf{m}} \hat{\varphi}_{\mathbf{n}} d\sigma &= \hat{\varphi}_{m_{p_C}} (x_{p_C}) \hat{\varphi}_{n_{p_C}} (x_{p_C}) h^2 \int_{\Gamma^C} \gamma \prod_{k=1, k \neq p_C}^3 \hat{\varphi}_{m_k} (x_k) \prod_{k=1, k \neq p_C}^3 \hat{\varphi}_{n_k} (x_k) \prod_{k=1, k \neq p_C}^3 d\sigma_k \\ &= \gamma \hat{\varphi}_{m_{p_C}} (x_{p_C}) \hat{\varphi}_{n_{p_C}} (x_{p_C}) \prod_{k=1, k \neq p_C}^3 h^2 \int_{[0,1]} \hat{\varphi}_{m_k} (x_k) \hat{\varphi}_{n_k} (x_k) dx_k \end{aligned}$$

which clearly implies that

$$\int_{\Gamma^C} \gamma \hat{\varphi}_{\mathbf{m}} \hat{\varphi}_{\mathbf{n}} d\sigma = \gamma \hat{\varphi}_{m_{p_C}} (x_{p_C}) \hat{\varphi}_{n_{p_C}} (x_{p_C}) \prod_{k=1, k \neq p_C}^3 M_{1,p} (m_k, n_k)$$

which ends the proof.  $\square$

Finally, using the two lemmas, we obtain

$$K_{3,p}(\mathbf{m}, \mathbf{n}) = \sum_{i=1}^3 \left( K_{1,p}(m_i, n_i) \prod_{k=1, k \neq i}^3 M_{1,p}(m_k, n_k) \right). \quad (38)$$

• *Proof of 3.*

To rewrite the terms  $K_{3,p}^C(\mathbf{m}, \mathbf{n})$  for all  $\mathbf{m}, \mathbf{n}$ , we use a similar reasoning as for  $K_{3,p}$ .

• *Proof of 4.*

To prove 4. and 5., we need the following lemma.

**Lemma C.4.** *Let  $\mathbf{m} = (m_p)_{p=1,\dots,3}$  and  $\mathbf{n} = (n_p)_{p=1,\dots,3}$ . We have*

$$M_{3,p}^{-1}(\mathbf{m}, \mathbf{n}) = \prod_{k=1}^3 M_{1,p}^{-1}(m_k, n_k).$$

*Proof.* Let  $A$  be the matrix defined  $\forall \mathbf{m}, \mathbf{n} \in \{1, \dots, p+1\}^d$  by

$$A(\mathbf{m}, \mathbf{n}) = \prod_{k=1}^3 M_{1,p}^{-1}(m_k, n_k).$$

We have

$$\begin{aligned} (AM)(\mathbf{m}, \mathbf{n}) &= \sum_{l_1, \dots, l_3=1}^{p+1} \left( \prod_{k=1}^3 M_{1,p}^{-1}(m_k, l_k) \prod_{k=1}^3 M_{1,p}(l_k, n_k) \right) \\ &= \left( \sum_{l_1, l_2=1}^{p+1} \left( \prod_{k=1}^2 M_{1,p}^{-1}(m_k, l_k) \prod_{k=1}^2 M_{1,p}(l_k, n_k) \right) \sum_{l_3=1}^3 M_{1,p}^{-1}(m_3, l_3) M_{1,p}(l_3, n_3) \right) \\ &= \prod_{k=1}^3 \left( \sum_{l_k=1}^{p+1} M_{1,p}^{-1}(m_k, l_k) M_{1,p}(l_k, n_k) \right). \end{aligned}$$

But,

$$\sum_{l_k=1}^{p+1} M_{1,p}^{-1}(m_k, l_k) M_{1,p}(l_k, n_k) = (M_{1,p}^{-1} M_{1,p})(m_k, n_k) = \delta_{m_k, n_k}$$

so that

$$(AM)(\mathbf{m}, \mathbf{n}) = \prod_{k=1}^3 \delta_{m_k, n_k} = I(\mathbf{m}, \mathbf{n})$$

where  $I$  is the identity matrix, which ends the proof.  $\square$

Let us now consider the matrix  $N_{3,p} = M_{3,p}^{-1} K_{3,p}$ . First, we rewrite  $K_{3,p}$  as

$$K_{3,p} = \sum_{q=1}^3 T_q,$$

with  $T_q(\mathbf{m}, \mathbf{n}) = K_{1,p}(m_q, n_q) \prod_{k=1, k \neq q}^3 M_{1,p}(m_k, n_k)$ .

Then,  $M_{3,p}^{-1} K_{3,p} = \sum_{q=1}^3 M_{3,p}^{-1} T_q$  and, using lemma C.4

$$\begin{aligned} (M_{3,p}^{-1} T_1)(\mathbf{m}, \mathbf{n}) &= \sum_{l_1, \dots, l_3=1}^{p+1} \left( \prod_{k=1}^3 M_{1,p}^{-1}(m_k, l_k) K_{1,p}(l_1, n_1) \prod_{k=2}^3 M_{1,p}(l_k, n_k) \right) \\ &= \sum_{l_1, \dots, l_3=1}^{p+1} M_{1,p}^{-1}(m_1, l_1) K_{1,p}(l_1, n_1) M_{1,p}^{-1}(m_2, l_2) M_{1,p}(l_2, n_2) M_{1,p}^{-1}(m_3, l_3) M_{1,p}(l_3, n_3) \\ &= \sum_{l_1=1}^{p+1} M_{1,p}^{-1}(m_1, l_1) K_{1,p}(l_1, n_1) \prod_{k=2}^3 \left( \sum_{l_k=1}^{p+1} M_{1,p}^{-1}(m_k, l_k) M_{1,p}(l_k, n_k) \right). \end{aligned}$$

So that

$$(M_{3,p}^{-1} T_1)(\mathbf{m}, \mathbf{n}) = N_{1,p}(m_1, n_1) \delta_{m_2, n_2} \delta_{m_3, n_3}.$$

Performing the same calculations for  $T_2$  and  $T_3$ , we obtain

$$N_{3,p}(\mathbf{m}, \mathbf{n}) = \sum_{p=1}^3 N_{1,p}(m_p, n_p) \prod_{k=1, k \neq p}^3 \delta_{m_k, n_k}.$$

• *Proof of 5.*

We apply the technique used to prove 4. to show that

$$N_{3,p}^C(\mathbf{m}, \mathbf{n}) = N_{1,p}^W(m_{p_C}, n_{p_C}) \prod_{k=1, k \neq p_C}^3 \delta_{m_k, n_k}$$

□

## CONTENTS

1. Introduction	1
2. Interior Penalty Discontinuous Galerkin Discretization of the wave equation	2
3. Stability analysis	4
4. Study in the 1-Dimensional case	6
4.1. Fourier Analysis of the IPDG scheme in 1D	7
4.2. Study of the condition $\lambda_{\min} \geq 0$	8
4.3. The CFL condition	11
5. The $d$ -dimensional case	13
5.1. From the 3 dimensional case to the one dimensional case	15
5.2. Consequences on the stability analysis	16
5.3. Extension to rectangular or parallelepiped mesh	17
6. Numerical results	18
6.1. Behaviour of the CFL condition with respect to $\alpha$	18
6.2. Comparison with numerical experiments	19
7. Conclusion	20
Appendix A. Expression of the polynomial $p_\alpha$	21
Appendix B. Definition of $Q_{p,\alpha}$ and $\tilde{Q}_{p,\alpha}$	22

Appendix C. Proof of theorem 5.1	25
References	30

## REFERENCES

- [1] C. Agut and J. Diaz. Stability analysis of the interior penalty discontinuous galerkin method for the wave equation. *INRIA Research Report*, 2010.
- [2] M. Ainsworth, P. Monk, and W. Muniz. Dispersive and dissipative properties of discontinuous galerkin finite element methods for the second-order wave equation. *Journal of Scientific Computing*, 27, 2006.
- [3] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19, No. 4:742–760, 1982.
- [4] D. N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Unified analysis of discontinuous galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2002.
- [5] C. Baldassari. Modélisation et simulation numérique pour la migration terrestre par équation d’ondes. *PhD Thesis*, 2009.
- [6] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier-stokes equations. *J. Comput. Phys.*, 131:267–279, 1997.
- [7] G. Benitez Alvarez, A.F. Dourado Loula, E.G. Dutra do Carmo, and A. Alves Rochinha. A discontinuous finite element formulation for helmholtz equation. *Comput. Methods. Appl. Mech. Engrg.*, 195:4018–4035, 2006.
- [8] G. Cohen. *Higher-Order Numerical Methods for Transient Wave Equations*. Springer, Berlin, 2001.
- [9] S. Cohen, P. Joly, J.E. Roberts, and N. Tordjman. Higher-order triangular finite elements with mass-lumping for the wave equation. *SIAM J. Numer. Anal.*, Vol. 44, 6:2408–2431, 2006.
- [10] S. Cohen, P. Joly, and N. Tordjman. Higher-order finite elements with mass-lumping for the 1d wave equation. *Finite Elements in Analysis and Design*, Vol. 16, Issues 3-4:329–336, 1994.
- [11] J. D. De Basabe and M. K. Sen. Stability of the high-order finite elements for acoustic or elastic wave propagation with high-order time stepping. *Geophys. J. Int.*, 181:577–590, 2010.
- [12] S. Fauqueux. Eléments finis mixtes spectraux et couches absorbantes parfaitement adaptées pour la propagation d’ondes élastiques en régime transitoire. *PhD Thesis*, 2003.
- [13] M. J. Grote, A. Schneebeli, and D. Schötzau. Discontinuous galerkin finite element method for the wave equation. *SIAM J. on Numerical Analysis*, 44:2408–2431, 2006.
- [14] M. J. Grote and D. Schötzau. Convergence analysis of a fully discrete discontinuous galerkin method for the wave equation. *Preprint No. 2008-04*, 2008.
- [15] D. Komatitsch and J. Tromp. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys. J. Int.*, 139:806–822, 1999.
- [16] G. Seriani and E. Priolo. Spectral element method for acoustic wave simulation in heterogeneous media. *Finite Elements in Analysis and Design*, Vol. 16 Issues 3-4:337–348, 1994.
- [17] K. Shahbazi. An explicit expression for the penalty parameter of the interior penalty method. *J. of Computational Physics*, 205:401–407, 2005.
- [18] T. Warburton and J. S. Hesthaven. On the constants in *hp*-finite element trace inverse inequalities. *Comput. Methods Appl. Mech. Engrg.*, 192:2765–2773, 2003.